UNIVERSITAT JAUME·I

MATHEMATICS DEPARTMENT

MODELLING, ESTIMATION AND APPLICATIONS OF SECOND-ORDER SPATIO-TEMPORAL CHARACTERISTICS OF POINT PROCESSES

BY	DIRECTOR
Francisco Javier Rodríguez Cortés	*PhD.* **Jorge Mateu**

GRADE 2013-2014

A Dissertation Submitted as **Ph.D. Thesis** into the program **Ph.D. IN COMPUTATIONAL MATHEMATICS** for the Department of Mathematics of School of Technology and Experimental Sciences of the Jaume I University in Partial Fulfillment of the requirements for **Ph.D. Degree** by **Royal Ordinance 1393/2007** of 29 October of 2007.

Castellón, 17 of February of 2014.

Signature   Signature (Director)

Acknowledgements

To MMMMM, my women.

There is no branch of mathematics, however abstract, which may not some day be applied to phenomena of the real world.

- Lobatchevsky, Nikolai

[The universe] cannot be read until we have learnt the language and become familiar with the characters in which it is written. It is written in mathematical language, and the letters are triangles, circles and other geometrical figures, without which means it is humanly impossible to comprehend a single word.

- Galilei, Galileo

Abstract

There is an extensive literature on the analysis of point process data in both time and space, separately. However, methods for the analysis of spatio-temporal point processes are less well established. Many spatial processes of scientific interest also have a temporal component that may need to be considered when modelling the underlying phenomenon. The spatio-temporal behaviour analysis is fundamental in areas such as environmental sciences, climate prediction and meteorology, epidemiology, image analysis, agriculture, seismology and astronomy, and so spatio-temporal point processes, rather than purely spatial point processes, must then be considered as potential models. A natural starting point for the analysis of spatio-temporal point process data is to investigate the nature of any stochastic interactions among the points of the process. For these processes, second-order properties play an important role for exploratory and inferential analysis. Second-order methods provide indeed a natural starting point for such analysis.

This thesis is mainly focused on developing properties and estimators for second-order characteristics of spatio-temporal point processes, and every chapter adds some valuable information over the previous ones.

In Chapter 1 we present a theoretical framework of spatial and spatio-temporal point processes as a mathematical tool for dealing with the concepts shown along the next chapters of this thesis. The final part of this chapter consists of a first compilation of the most recent developments in the literature of spatio-temporal point processes.

In Chapter 2 we consider kernel-based non-parametric estimation of second-order product densities of spatial point patterns. We present a new family of optimal and positive kernels showing less variance than that for optimal kernels. This family generalises most of the classical and widely used kernel functions, such as Box or Epanechnikov kernels. We propose an alternative unbiased estimator for the product density function, and compare the performance of the estimator for several members of the family of optimal and positive kernels through MISE and relative efficiency.

We present a simulation study to analyse the behaviour of such kernel functions, for three different spatial structures, for which we know the exact analytical form of the product density, and under small sample sizes. Some known datasets are revisited.

In Chapter 3 a new kernel estimator of the second-order product density function of a spatio-temporal point process with and without considering first- and second-order spatio-temporal separability is given. The spatio-temporal second-order product density function is of interest as can be used to discriminate amongst several spatio-temporal point structures. Further, the expectation and variance of this estimator are obtained. In addition, as we have developed close expressions for the variance under the Poisson case, we use them to generate the corresponding confidence surfaces. A simulation study is presented. We have used functions of the R library stpp in connection with Fortran subroutines. Finally, we apply the resulting estimator to data on the spatio-temporal distribution of invasive meningococcal disease in Germany.

In Chapter 4 we focus on second-order orientation methods which provide a natural tool for the analysis of anisotropic spatial point process data. Here we extend to the spatio-temporal setting the spatial point pair orientation distribution function. The new spatio-temporal orientation distribution function is used to detect spatio-temporal anisotropic configurations. An edge-corrected estimator is defined and illustrated through a simulation study. We apply the resulting estimator to data on the spatio-temporal distribution of fire ignition events caused by humans in a square area of 30×30 km^2 during four years. Our results confirm that our approach is able to detect directional components at distinct spatio-temporal scales.

Finally, we provide a general description of the currently ongoing research projects which have emerged motivated by the close relationship with the second-order properties of spatial and spatio-temporal point processes. In particular, we have adapted our methodology to spatio-temporal local clustering analysis, and to modelling orbital debris using a new and innovative adaptation over the sphere of the classical theory.

Contents

List of Figures

List of Tables

Principal Notation

$b(x,r)$	:	disc of radius r centered at x
$d(x)$	:	distance from point x to its nearest neighbour
$\mathbb{E}$	:	expectation
$\mathbb{E}^!$	:	expectation with respect to Palm distribution
$g(r)$	:	pair correlation function
$g^{ij}(r)$	:	cross pair correlation function
$G(t)$	:	point density d.f.
$J(r)$	:	J-function
$K(r)$	:	Ripley's K-function
$K^{ij}(r)$	:	multivariate K-function
$\mathcal{K}(B)$	:	second-order reduced moment measure
$L(r)$	:	L-function
$L^{ij}(r)$	:	multivariate L-function
X	:	point process
$N(A)$	:	number of points of X in A
o	:	origin of space $\mathbb{R}^d$
$\mathbb{P}$	:	probability
$\mathbb{P}^!_o$	:	probability with respect to palm distribution
r	:	distance variable
$\mathbb{R}$	:	real line
$\mathbb{R}^d$	:	d-dimensional Euclidean space
$\alpha^{(k)}$	:	k-th factorial moment measure

$\gamma_W(r)$:   isotropic set covariance

ρ :   intensity

$|W|$:   area of W

$|T|$:   length of T

$\rho^{(2)}(r)$:   second-order product density in stationary and and isotropic case

$\rho^{(k)}(r)$: k-th product density

$\sum^{\neq}$:   sum over point pairs or k-tuples of distinct point

$\mathbf{1}_A[x]$:   indicator function

$\| \cdot \|$:   Euclidean distance

2

Chapter 1

Point processes methodology

In order to provide a unified and proper context to develop our methodology, we consider in this chapter a set of definitions and important results that provide subsequently a nice grammatical and mathematical basis. We take Stoyan and Stoyan (1994), Møller and Waagepetersen (2004), Illian et al. (2008) and Chiu et al. (2013) texts as a fundamental references, treating deeply and rigorously all the concepts that we need in the rest of the work.

We begin recalling the definition and basic concepts of point processes in general metric spaces, we give a the theoretical description of first- and secon-order characteristics such as the intensity function, the pair correlation function, K-function, J-function and orientation distribution function. We also give the definition of certain types of models for spatial point patterns such as Cox processes, geometric anisotropic processes and some special cases of Markov processes following Ripley (1988), Diggle et al. (2000), Illian et al. (2008) and Møller and Waagepetersen (2004). We also give an extension of this theory to the case of multivariate point processes (Lotwick and Silverman (1981)). We present statistical estimations of these first- and second-order characteristics in the spatial context using non-parametric methods.

Finally we show a sophisticated approach to the definition of spatio-temporal first- and second-order characteristics taken from Diggle et al.

(1995), Møller and Diaz-Avalos (2010), Møller and Ghorbani (2010), Diggle and Gabriel (2010), Ghorbani (2013), Gabriel (2013), Diggle (2013) and Cronie and Lieshout (2014) and well as the corresponding properties and non-parametric estimators. This last part of the chapter is built as an attempt to gather all the recent literature on this methodology to build the path towards the context spatio-temporal point processes.

1.1 Point processes on metric spaces

Let S be a metric space with metric $d(\cdot, \cdot)$ and $\mathcal{B}$ be the Borel σ-algebra (generated by open sets) in S. Let $\mathcal{B}_0 \subseteq \mathcal{B}$ be the system of all bounded Borel sets. We define the space of locally finite subsets of S as

$$N_{\mathrm{lf}} = \{x \subseteq S : n(x_B) < +\infty, \ \forall B \in \mathcal{B}_0\},$$

where $x_B = x \cap B$ and $n(y)$ denotes the cardinality of the set y. Elements of N_{lf} are called *locally finite point configurations*. We equip N_{lf} with the σ-algebra

$$\mathcal{N}_{\mathrm{lf}} = \sigma\left\{\{x \in N_{\mathrm{lf}} : n(x_B) = m\}, m \in \mathbb{N}_0, B \in \mathcal{B}_0\right\},$$

where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. A point process is defined as a random locally finite point configuration.

Definition 1 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an abstract probability space. A point process X is a measurable mapping*

$$X : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (N_{\mathrm{lf}}, \mathcal{N}_{\mathrm{lf}}).$$

Definition 2 *The distribution of the simple point processs is a measure P_X defined on $(N_{lf}, \mathcal{N}_{lf})$ defined by the relation*

$$P_X(F) = \mathbb{P}(X \in F) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in F\}), \qquad F \in \mathcal{N}_{\mathrm{lf}}.$$

We say that the point process is finite if $n(X) < +\infty$ almost surely.

Definition 3 *For a point process X we will denote the number of points in the set B by $N(B) = n(X_B) = n(X \cap B)$ and refer to the function N as a count function.*

Proposition 4 *X is a point process if and only if $N(B)$ is a random variable for any $B \in \mathcal{B}_0$.*

Definition 5 *Let X be a point process. By void probabilities we understand probabilities $\mathbb{P}(N(B) = 0), B \in \mathcal{B}_0$.*

Theorem 6 *A point process is uniquely determined by its void probabilities.*

Definition 7 *A point process X is stationary if its distribution is translation invariant, i.e. $X + \mathbf{u} = \{\xi + \mathbf{u} : \xi \in X\}$ has the same distribution as X for any $\mathbf{u} \in S$. A point process X is isotropic if its distribution is invariant under rotations around the origin, i.e. $\mathcal{O}X = \{\mathcal{O}\xi : \xi \in X\}$ has the same distribution as X for any rotation $\mathcal{O}$ around the origin.*

Definition 8 *The intensity measure μ on S is given by*

$$\mu(B) = \mathbb{E}[N(B)], \qquad B \in \mathcal{B}.$$

Definition 9 *If the intensity measure μ can be written as*

$$\mu(B) = \int_B \rho(\xi)\,\mathrm{d}\xi, \qquad B \in \mathcal{B},$$

where ρ is a non-negative function, then ρ is called the intensity function. If ρ is constant, then X is said to be homogeneous or first-order stationary with intensity ρ; otherwise X is said to be inhomogeneous.

Heuristically, $\rho(\xi)\,\mathrm{d}\xi$ is the probability for the occurrence of a point in an infinitesimally small ball with centre ξ and volume $\mathrm{d}\xi$. For a homogeneous point process, ρ is the mean number of points per unit volume and the constant is called *intensity* or *rate*.

Proposition 10 *If X is a stationary point process with locally finite intensity measure μ, then μ is proportional to the Lebesgue measure.*

Since the intensity measure of a stationary point process is proportional to the Lebesgue measure, the intensity function is constant and equal to this proportionality constant. It means that every stationary point process is homogeneous.

1.2   Moment measures

The first- and higher-order moments of the counts $N(B)$ with $B \in \mathcal{B}$, can be expressed by the following measures.

Definition 11 *For a point process X on S and each $m \in \mathbb{N}$, define the m-th order moment measure $\mu^{(m)}$ on S^m by*

$$\mu^{(m)}(D) = \mathbb{E}\left[\sum_{\xi_1,\dots,\xi_m} \mathbf{1}\left[(\xi_1,\dots,\xi_m) \in D\right]\right], \qquad D \subseteq S^m,$$

and the m-th order order factorial moment measure $\alpha^{(m)}$ on S^m by

$$\alpha^{(m)}(D) = \mathbb{E}\left[\sum_{\xi_1,\dots,\xi_m}^{\neq} \mathbf{1}\left[(\xi_1,\dots,\xi_m) \in D\right]\right], \qquad D \subseteq S^m,$$

where the $\sum^{\neq}$ over the summation sign means that the m points $\xi_1,\dots,\xi_m$ are pairwise distinct and $\mathbf{1}[\cdot]$ denotes the indicator function.

In particular, $\mu = \mu^1 = \alpha^1$ is called the *intensity measure*. The m-th order moment measure $\mu^{(m)}$ determines the m-th order moments of the count variables $N(B)$, $B \subseteq S$, since

$$\mu^{(m)}(B_1 \times \cdots \times B_m) = \mathbb{E}\left[\prod_{i=1}^{m} N(B_i)\right], \qquad B_i \in \mathcal{B}.$$

For any $m \in \mathbb{N}$, there is a one-to-one correspondence between $(\mu^{(1)}, \ldots, \mu^{(m)})$ and $(\alpha^{(1)}, \ldots, \alpha^{(m)})$. It is often more convenient to work with the reduced moment measures.

The above definition immediately extends to

$$\mathbb{E}\left[\sum_{\xi_1,\ldots,\xi_m}^{\neq} h\left(\xi_1, \ldots, \xi_m\right)\right] = \int_{S^m} h\left(\xi_1, \ldots, \xi_m\right) \mathrm{d}\alpha^{(m)}\left(\xi_1, \ldots, \xi_m\right), \quad (1.1)$$

for non-negative functions $h : S^m \times N_{\mathrm{lf}} \longrightarrow [0, +\infty)$.

1.2.1   The second-order reduced moment measure

Notation 12 *Let $\mathcal{B}^d$ be the Borel σ-algebra (generated by open sets) in $\mathbb{R}^d$.*

Definition 13 *If the second-order factorial moment measure $\alpha^{(2)}$ can be written as*

$$\alpha^{(2)}\left(C\right) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbf{1}\left[(\xi, \eta) \in C\right] \rho^{(2)}\left(\xi, \eta\right) \mathrm{d}\xi \mathrm{d}\eta, \qquad C \subseteq \mathbb{R}^d \times \mathbb{R}^d,$$

where $\rho^{(2)}$ is a non-negative function, then $\rho^{(2)}$ is called the second-order product density function.

Definition 14 *If both ρ and $\rho^{(2)}$ exist, the pair correlation function is defined by*

$$g(\xi, \eta) = \frac{\rho^{(2)}(\rho, \eta)}{\rho(\xi)\rho(\eta)}, \qquad (1.2)$$

where we take $a/0 = 0$ for $a \geq 0$.

Definition 15 *Suppose that X has intensity function ρ and that the measure*

$$\mathcal{K}(B) = \frac{1}{|A|}\mathbb{E}\left[\sum_{\xi,\eta \in X}^{\neq} \frac{\mathbf{1}\left[\xi \in A, \eta - \xi \in B\right]}{\rho\left(\xi\right)\rho\left(\eta\right)}\right], \qquad B \subseteq \mathbb{R}^d \qquad (1.3)$$

7

does not depend on the choice of $A \subseteq \mathbb{R}^d$ with $0 < |A| < +\infty$, where we take $a/0 = 0$ for $a \geq 0$. Then X is said to be second-order intensity reweighted stationary and $\mathcal{K}$ is called the second-order reduced moment measure.

Proposition 16 *If the pair correlation function exists and is invariant under translations, then we have second-order intensity reweighted stationarity and*

$$\mathcal{K}(B) = \int_B g(\xi)\mathrm{d}\xi, \qquad B \subseteq \mathbb{R}^d. \tag{1.4}$$

1.2.2 Campbell measures and Palm distributions

Assume that μ is σ-finite, i.e. $\mu(B_i) < +\infty$ for a countable partition B_i of S (this is e.g. satisfied if μ is locally finite).

Definition 17 *The Campbell measure is defined by*

$$C(D) = \mathbb{E}\left[\sum_{\xi \in X} \mathbf{1}\left[(\xi, X) \in D\right]\right], \qquad D \in \mathcal{B} \times \mathcal{N}_{\mathrm{lf}}.$$

Definition 18 *For a point process X on S, define the reduced Campbell measure $C^!$ on $S \times N_{\mathrm{lf}}$ by*

$$C^!(D) = \mathbb{E}\left[\sum_{\xi \in X} \mathbf{1}\left[(\xi, X \backslash \{\xi\}) \in D\right]\right], \qquad D \in \mathcal{B} \times \mathcal{N}_{\mathrm{lf}}.$$

Hence the term reduced Campbell measure for $C^!$. We have

$$\mathbb{E}\left[\sum_{\xi \in X} \mathbf{1}\left[(\xi, X \backslash \{\xi\}) \in D\right]\right] = \int h(\xi, x)\,\mathrm{d}C^!(\xi, x),$$

for non-negative functions h. Note that $C^!$ determines (μ, α^2), since clearly $\mu(\cdot) = C^!(\cdot \times N_{\mathrm{lf}})$ and

$$\alpha^{(2)}(B_1 \times B_2) = \int \mathbf{1}\left[\xi \in B_1\right] n\left(x_{B_2}\right)\mathrm{d}C^!(\xi, x).$$

8

For each $F \in \mathcal{N}_{\mathrm{lf}}$, $C^!\left(\cdot \times F\right) \leq \mu\left(\cdot\right)$, so $C^!\left(\cdot \times F\right)$ is absolutely continuous $\left(C^!\left(\cdot \times F\right) \ll \mu\right)$ with respect to μ. Then by the Radon-Nikodym theorem, there exists a μ-almost surely unique integrable function $P_\xi^!$ such that

$$C^!\left(B \times F\right) = \int_B P_\xi^!\left(F\right) \mathrm{d}\mu\left(\xi\right), \qquad (1.5)$$

where $P_\xi^!\left(\cdot\right)$ is a probability measure for each $\xi \in S$; for details, (see e.g. Daley and Vere-Jones (2008)).

Definition 19 *The probability measure $P_\xi^!\left(\cdot\right)$ on $\mathcal{N}_{\mathrm{lf}}$ is called a reduced Palm distribution at point ξ.*

From (1.5) we obtain the so called *Campbell-Mecke Theorem.*

Theorem 20 (The Campbell-Mecke theorem)

$$\mathbb{E}\left[\sum_{\xi \in X} h\left(\xi, X \backslash \{\xi\}\right)\right] = \iint h\left(\xi, x\right) \mathrm{d}P_\xi^!\left(x\right) \mathrm{d}\mu\left(\xi\right), \qquad (1.6)$$

for non-negative functions h.

Theorem 21 (Slivnyak-Mecke's theorem) *If $X \sim Poisson(S, \rho)$, then for functions $h : S \times N_{\mathrm{lf}} \to [0, \infty)$,*

$$\mathbb{E}\left[\sum_{\xi \in X} h\left(\xi, X \backslash \{\xi\}\right)\right] = \int_S \mathbb{E}\left[h\left(\xi, x\right)\rho(\xi)\right] \mathrm{d}\xi. \qquad (1.7)$$

Assume that X is a stationary point process on $\mathbb{R}^d$ with intensity $0 < \rho < +\infty$. For $\xi \in \mathbb{R}^d$, $x \in N_{\mathrm{lf}}$, y $F \subseteq N_{\mathrm{lf}}$, let $x + \xi = \{\eta + \xi : \eta \in x\}$ denote the translation of the point configuration x by ξ, and $F + \xi = \{x + \xi : x \in F\}$ the translation of F by ξ.

Theorem 22 *In the stationary case*

$$P_0^!\left(F\right) = \mathbb{E}\left[\sum_{\xi \in X_B} \frac{\mathbf{1}\left[X \backslash \{\xi\} \in F + \xi\right]}{\rho \left|B\right|}\right], \qquad F \subseteq N_{\mathrm{lf}}, \qquad (1.8)$$

9

for an arbitrary set $B \subseteq S$ with $0 < |B| < +\infty$, and $P_\xi^! (F) = P_0^! (F - \xi)$.
Moreover

$$\mathbb{E}\left[\sum_{\xi \in X} h\left(\xi, X \backslash \{\xi\}\right)\right] = \rho \iint h\left(\xi, x + \xi\right) \mathrm{d}P_0^!\left(x\right) \mathrm{d}\xi, \qquad (1.9)$$

for non-negative functions h.

Consider the problem of estimating $P_0^!(F)$ for some $F \in \mathcal{B}$. Since X is stationary we may consider $X - \xi$, $\xi \in X$, for a bounded B. as representing observations of X conditional on $0 \in X$. Thus a natural estimator of $P_0^!(F)$ is the empirical average

$$\frac{1}{N(B)} \sum_{\xi \in X_B} \mathbf{1}\left[X \backslash \{\xi\} \in F + \xi\right].$$

If B is large, we may expect that $\rho \approx N(B) / |B|$, and so we obtain the estimator

$$\frac{1}{\rho |B|} \sum_{\xi \in X_B} \mathbf{1}\left[X \backslash \{\xi\} \in F + \xi\right].$$

This is exactly the unbiased estimator of $P_0^!(F)$ obtained from (1.8).

1.2.3 Interpretation of $\mathcal{K}$ as Palm expectation

The second-order reduced moment measure $\mathcal{K}$ has an interpretation as a Palm expectation, since

$$\mathcal{K}(B) = \int_{\mathbb{R}^d} \sum_{\eta \in X_B} \frac{\mathbf{1}\left[\eta - \xi \in B\right]}{\rho(\eta)} \mathrm{d}P_\xi^!(x), \qquad (1.10)$$

for almost all $\xi \in \mathbb{R}^d$. This follows from the Campbell-Mecke theorem (1.6). In the stationary case, by (1.9),

$$\mathcal{K}(B) = \frac{\mathbb{E}_0^!\left(N(B)\right)}{\rho}, \qquad B \subseteq \mathbb{R}^d,$$

where $E_0^!$ denotes expectation with respect to $P_0^!$.

10

1.3 Models

1.3.1 Poisson point process

Poisson point processes play a fundamental role. They serve as a tractable model class for *no interaction* or *complete spatial randomness* in spatial point patterns. They also serve as reference processes when summary statistics are studied.

Definition 23 *Let μ be a diffuse measure (i.e. $\mu(\{\xi\}) = 0, \forall \xi \in S$), $n \in \mathbb{N}$ and $B \in \mathcal{B}$ such that $0 < \mu(B) < +\infty$. Consider $X = (\xi_i)_{i=1}^n$ independent random vectors with identical distribution f. X is called binomial point process of n points in B with density f. We write $X \sim Binomial(B, n, f)$.*

Definition 24 *A point process X on S is a Poisson point process with intensity function ρ if the following properties are satisfied (where μ is given by* **Definition 23***):*

   i. For any $B \in \mathcal{B}$ with $\mu(B) < +\infty$, $N(B) \sim Poisson(\mu(B))$ a Poisson distribution with mean $\mu(B)$.

   ii. For any $n \in \mathbb{N}$ and $B \in \mathcal{B}$ with $0 < \mu(B) < +\infty$, conditional on $N(B) = n$, $X_B \sim Binomial(B, n, f)$ with $f(\xi) = \rho(\xi)/\mu(\xi)$.

We write $X \sim Poisson(S, \rho)$.

Definition 25 *If ρ is constant, the process $Poisson(S, \rho)$ is called a homogeneous Poisson process on S with rate or intensity ρ; else it is said to be an inhomogeneous Poisson process on S.*

Proposition 26 *A homogeneous Poisson point process is a stationary and isotropic process.*

Theorem 27 *A Poisson point process exists and it is uniquely determined by its intensity measure.*

Theorem 28 *If $X_i \sim Poisson(S, \rho_i)$, for $i \in \mathbb{N}$ mutually independent and*

$$\rho = \sum_{i=1}^{+\infty} \rho_i,$$

is locally integrable, then with probably one,

$$X = \bigcup_{i=1}^{+\infty} X_i,$$

is a disjoint union, and $X \sim Poisson(S, \rho)$.

1.3.2　Cox process

A Cox process is a natural extension of a Poisson process, obtained by considering the intensity function of the Poisson process as non-constant and as a realisation of a random field (Adler (1981)). Such processes were studied in a seminal paper by Cox (1955) under the name *doubly stochastic Poisson processes*, but are today usually called Cox processes. We begin by assuming that $S \subseteq \mathbb{R}^d$.

Definition 29 *Z is a random field if $Z(\xi)$ is a random variable for all $\xi \in S$. If $\rho(\xi) = \mathbb{E}(Z(\xi))$ exists and is locally integrable, then with probability one, $Z(\xi)$ is a locally integrable function. The intensity measure of the Poisson process $X|Z$ is*

$$\mu(B) = \int_B Z(\xi)\, \mathrm{d}\xi, \qquad B \in \mathcal{B}. \tag{1.11}$$

Definition 30 *Suppose that $Z = \{Z(\xi) : \xi \in S\}$ is a non-negative random field so that with probability one, Z is a locally integrable function. If the conditional distribution of X given Z is a Poisson process on S with intensity function Z, then X is said to be a Cox process driven by Z.*

Theorem 31 *Consider a Cox process X, then $X|Z$ is a Poisson process. We obtain the following results:*

i. The void probabilities are given by

$$v(B) = \mathbb{E}\left[\exp\left(-\int_B Z(\xi)\,\mathrm{d}\xi\right)\right], \qquad B \in \mathcal{B}.$$

ii. The intensity function is

$$\rho(\xi) = \mathbb{E}[Z(\xi)].$$

1.3.3 The Neyman-Scott process

We consider here Neyman-Scott processes, as proposed by Neyman and Scott (1958) which are also particular Cox processes.

Definition 32 *Let C be a stationary Poisson process on $\mathbb{R}^d$ with intensity $\kappa > 0$. Conditional on C, let $X_c, c \in C$, be independent Poisson processes on $\mathbb{R}^d$ where X_c has intensity function*

$$\rho_c(\xi) = \alpha\kappa(\xi - c),$$

where $\alpha > 0$ is a parameter and κ is a kernel (i.e. for all $c \in \mathbb{R}^d$, $\kappa(\xi - c)$ is a density function). Then

$$X = \bigcup_{c \in C} X_c,$$

is a special case of a Neyman-Scott process with cluster centres C and clusters $X_c, c \in C$ (in the general definition of a Neyman-Scott process, $n(X_c)$ given C is not restricted to be a Poisson variate, see e.g. Stoyan et al. (1995). By **Theorem 28***, X is also a Cox process on $\mathbb{R}^d$ driven by*

$$Z(\xi) = \sum_{c \in C} \alpha\kappa(\xi - c). \tag{1.12}$$

Clearly, Z in (1.12) is stationary and locally integrable, and it is also isotropic if $\kappa(\xi) = \kappa(\|\xi\|)$ is isotropic. The intensity is $\rho = \alpha\kappa$, and the

pair correlation function is given by

$$g(\xi) = 1 + h(\xi)/k,$$

where

$$h(\xi) = \int k(\eta)k(\xi + \eta)\mathrm{d}\eta, \qquad (1.13)$$

is the density for the difference between two independent points, each having density k.

1.3.4 The hard-core process

Definition 33 *Let ω_d and σ_d be the volume and the surface area of the d-dimensional unit ball, respectively. Then*

$$\omega_d = \frac{\pi^{d/2}}{\Gamma(1 + d/2)}, \quad \sigma_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}, \qquad (1.14)$$

where Γ is the Gamma function, (Whittaker and Watson (1996)).

Under the definition of the Neyman-Scott process, the hard-core repulsion process or Matérn process, Matérn (1986) is given by

$$\kappa(\xi) = \frac{\mathbf{1}\left[\|\xi\| \le r\right]}{\omega_d r^d},$$

the uniform density on the ball $b(\mathbf{0}, r)$.

Another construction of a hard-core process is based on the sequential approach.

Definition 34 *Let $r > 0$ and $B \in \mathcal{B}_0$ be given. A simple sequential inhibition (SSI) process in the set B is constructed in the following way:*

  i. choose $\xi_1 \in B$ uniformly at random,

  ii. if $k - 1$ points are chosen, choose ξ_k uniformly in $B \setminus \bigcup\limits_{i=1}^{k-1} b(\xi_i, r)$,

  iii. the construction ends in n steps, if $B \subseteq \bigcup\limits_{i=1}^{n} b(\xi_i, r)$.

14

Since $B \backslash \bigcup\limits_{i=1}^{k-1} b(\xi_i, r)$ can have complicated geometrical shape, in practice the process is simulated by the rejection method. A point ξ_i is generated uniformly in the window B and if it lies closer than r from an existing point, then it is rejected and a new point is generated.

1.3.5   Geometric anisotropic processes

Møller and Toftaker (2012) consider spatial point processes with a pair correlation function $g(\mathbf{u})$ which depends only on the lag vector $\mathbf{u} \in \mathbb{R}^2$ between a pair of points. Their interest is in statistical models with a special kind of *structured* anisotropy: g is geometric anisotropy if it is elliptical but not spherical.

Definition 35 *Let g elliptical, i.e.*

$$g(\mathbf{u}) = g_0(\sqrt{\mathbf{u}\Sigma^{-1}\mathbf{u}^t}), \qquad \mathbf{u} \in \mathbb{R}^2$$

where u is a row vector wiht transpose u^t, the function g_0 is such that g is locally integrable. Σ is a 2×2 symmetric positive definite matrix of the form

$$\Sigma = \omega^2 U_\theta \operatorname{diag}(1, \zeta^2) U_\theta^t,$$

where $0 \leq \theta < \pi$, $\omega > 0$, with $0 < \zeta \leq 1$ is the ratio of the minor axis and the major axis; we call it the anisotropy factor. Finally U_θ is the orthonormal matrix

$$U_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

The ellipse $E = \{\mathbf{u} : \mathbf{u}\Sigma^{-1}\mathbf{u}^t = 1\}$ as semi-major axis ω corresponding to the angle θ, and semi-minor axis $\omega\xi$ corresponding to the angle $\theta + \pi/2$. In the isotropic case of g, we have that $\xi = 1$, $\Sigma = \omega^2 I$, E is a circle of radius ω, and the value of θ plays no role.

Geometric anisotropic Cox processes

Assume that $Z = \{Z(\mathbf{u}) : \mathbf{u} \in \mathbb{R}^2\}$ is a non-negative second-order stationary random field with mean one and covariance function

$$c(\mathbf{u}) = \mathbb{E}[Z(\mathbf{0})Z(\mathbf{u})] - 1, \qquad \mathbf{u} \in \mathbb{R}^2,$$

and that X conditional on Z is a Poisson process with an intensity function of the multiplicative form

$$\lambda(\mathbf{u}) = \rho(\mathbf{u})Z(\mathbf{u}), \qquad \mathbf{u} \in \mathbb{R}^2,$$

then X is a Cox process driven by (Cox (1955)).

Assuming X has intensity function ρ, and X is second-order intensity-reweighted stationary, the pair correlation function is of the form

$$g(\mathbf{u}) = 1 + c(\mathbf{u}) = \mathbb{E}[Z(\mathbf{0})Z(\mathbf{u})], \qquad \mathbf{u} \in \mathbb{R}^2.$$

The residual function of the processes is given by

$$k_\nu(r) = \frac{r^\nu K_\nu(r)}{\pi 2^{\nu+1}\Gamma(\nu+1)}, \qquad r \geq 0,$$

where $\nu > -1/2$ and K_ν is the modified Bessel function of the second kind. The functions $\mathbf{u} \to k_n(\|\mathbf{u}\|/\omega)/\kappa$ with $\nu > 0$, $\omega > 0$, and $\kappa > 0$ provide a flexible class of isotropic covariance functions known as Whittle-Matérn covariance function.

Geometric anisotropic Log Gaussian Cox processes

Assuming $\log Z$ is a stationary Gaussian random field, then X is a log Gaussian Cox process (LGCP). Note that X is stationary if and only if ρ is constant. Denoting c the covariance function of $\log Z$, the assumption that $\mathbb{E}[\log Z] = 1$ means that $\mathbb{E}[\log Z] = -c(\mathbf{0})/2$.

Since

$$g(\mathbf{u}) = \exp(c(\mathbf{u})), \qquad \mathbf{u} \in \mathbb{R}^2,$$

is we assume that c is elliptical, i.e.

$$c(\mathbf{u}) = c_0(\sqrt{\mathbf{u}\Sigma^{-1}\mathbf{u}^t}), \qquad \mathbf{u} \in \mathbb{R}^2,$$

when

$$c_0(r) = k_\nu(r)/\kappa, \qquad r \geq 0,$$

where $\nu > 0$ and $\kappa > 0$ are parameters, then the pair correlation take the form

$$g(\mathbf{u}) = \exp\left(k_\nu\left(\sqrt{\mathbf{u}\Sigma^{-1}\mathbf{u}^t}\right)/\kappa\right), \qquad \mathbf{u} \in \mathbb{R}^2.$$

We refer to this model for X as the Whittle-Matérn LGCP. Simulation of Gaussian random fields and LGCPs is discussed in Møller and Waagepetersen (2004) and the references therein.

Geometric anisotropic Shot noise Cox processes

Let X be a stationary Poisson process on $\mathbb{R}^2$ with intensity $\kappa > 0$, and f be a quadratically integrable density function on $\mathbb{R}^2$. Define the residual process by

$$R(\mathbf{u}) = \frac{1}{\kappa} \sum_{\mathbf{v} \in \mathbb{R}^2} f(\mathbf{u} - \mathbf{v}), \qquad \mathbf{u} \in \mathbb{R}^2.$$

Clearly, R is stationary and $\mathbb{E}[R(\mathbf{u})] = 1$, while X is stationary if and only if ρ is constant. Assume that $f = f_\Sigma$ is elliptical, i.e.

$$f_\Sigma(\mathbf{u}) = f_0\left(\sqrt{\mathbf{u}\Sigma^{-1}\mathbf{u}^t}\right)|\Sigma|^{-1/2}, \qquad \mathbf{u} \in \mathbb{R}^2.$$

If the covariance is defined as

$$c(\mathbf{u}) = f_\Sigma * f_\Sigma/\kappa = f_I * f_I(\mathbf{u}\Sigma^{-1})/[\kappa|\Sigma|^{-1/2}], \qquad \mathbf{u} \in \mathbb{R}^2.$$

where $*$ denotes convolution, then it is elliptical. Taking $f_0 = k_\nu$ with $\nu > -1/2$, then the pair correlation is of the form

$$g(\mathbf{u}) = 1 + k_{2\nu+1}\left(\sqrt{\mathbf{u}\Sigma^{-1}\mathbf{u}^t}\right)/[\kappa|\Sigma|^{-1/2}], \qquad \mathbf{u} \in \mathbb{R}^2. \qquad (1.15)$$

1.4   Summary statistics

1.4.1   Second-order summary statistics

Definition 36 *The K- and L-functions for a second-order reweighted stationary point process are defined by*

$$K\left(r\right) = \mathcal{K}\left(b\left(\mathbf{0}, r\right)\right), \qquad L\left(r\right) = \left(\frac{K\left(r\right)}{\omega_d}\right)^{1/d}, \qquad r > 0. \qquad (1.16)$$

This definition, which extends the definition of Ripley's K-function (Ripley (1976, 1977)) for the stationary case to the case of second-order intensity reweighted stationarity, is due to Baddeley et al. (2000). In the stationary case, $\rho K\left(r\right)$ is the expected number of further points within distance r from the origin given that X has a point at the origin.

Proposition 37 *For a stationary and isotropic process*

$$K(r) = \sigma_d \int_0^r u^{d-1} g(u)\mathrm{d}u, \quad or \quad g(r) = \frac{1}{\sigma_d r^d} \frac{\mathrm{d}K(r)}{\mathrm{d}r}. \qquad (1.17)$$

Theorem 38 *For a stationary Poisson point process*

$$K\left(r\right) = \omega_d r^d = \frac{\pi^{d/2} r^d}{\Gamma\left(1 + \frac{d}{2}\right)} \qquad and \qquad L\left(r\right) = r. \qquad (1.18)$$

So, in the particular case of $d = 2$,

$$K\left(r\right) = \pi r^2 \qquad and \qquad g(r) = 1. \qquad (1.19)$$

The K- and L-functions are in one-to-one correspondence, and in applications the L-function is often used instead of the K-function. One reason is that L is the identity for a Poisson process. In general, at least for small values of r, $L(r) - r > 0$ indicates aggregation or clustering at distances less than r, and $L(r) - r < 0$ regularity at distances less than r. This may be due to certain latent processes or attraction or repulsion between the

points. Moreover, for a homogeneous Poisson process, the transformation $K \to L$ is variance stabilising when K is estimated by non-parametric methods (Besag (1977)).

1.4.2 Summary statistics based on inter-point distances

Definition 39 *Assume that X is stationary. The* empty space function *F is the distribution function of the distance from the origin (or another fixed point $\mathbb{R}^d$) to the nearest point in X, i.e.*

$$F(r) = \mathbb{P}(X \cap b(\mathbf{0}, r) \neq \emptyset), \qquad r > 0. \tag{1.20}$$

The *nearest-neighbour function G* is

$$G(r) = \frac{1}{\rho|A|}\mathbb{E}\left[\sum_{\xi \in X \cap A} \mathbf{1}[(X \backslash \{\xi\}) \cap b(\xi, r) \neq \emptyset]\right], \qquad r > 0, \tag{1.21}$$

for an arbitrary se $A \subset \mathbb{R}^d$ wiht $0 < |A| < +\infty$. For the nearest-neighbour function G in (1.21) by **Theorem 20** it follows that

$$G(r) = \mathbb{P}_0^!(N(b(\mathbf{0}, r)) > 0), \qquad r > 0.$$

The *J*-function is defined by

$$J(r) = \frac{1 - G(r)}{1 - F(r)}, \quad \text{for} \quad F(r) < 1.$$

The *J*-function was suggested by Lieshout and Baddeley (1996).

Remark 40 *By stationarity, (1.21) does not depend on the choice of A.*

Lemma 41 *For a stationary Poisson process on $\mathbb{R}^d$ with intensity $\rho < \infty$,*

$$F(r) = G(r) = 1 - \exp(-\rho\omega_d r^d) \quad and \quad J(r) = 1 \quad r > 0,$$

where the equality for G coming from the Slivnyak-Mecke theorem **Theorem 21**.

Remark 42 *In general, at least for small values of $r > 0$, $F(r) > G(r)$ (or $J(r) < 1$) indicates clustering, and $F(r) > G(r)$ (or $J(r) > 1$) indicates regularity, but if $J(r) = 1$ does not imply that X is a stationary Poisson process, (see Bedford and van den Berg (1997)).*

Lemma 43 *For Neyman-Scott point process the J-function can be expressed as*

$$J(r) = \int k(\xi) \exp \left(-\alpha \int\limits_{\|\eta\| \leq r} k(\xi + \eta) \mathrm{d}\eta \right) \mathrm{d}\xi,$$

where k is given by (1.13), thus $J(r)$ is non-increasing for $r > 0$ with range $(\exp(-\alpha), 1)$. So $F(r) < G(r)$ for $r > 0$.

1.4.3 Summary directional statistics

In order to investigate for a possible *anisotropy* or *directionality* in the planar case, an orientation methodology was considered by Stoyan and Stoyan (1994), Brix and Møller (2001) and Illian et al. (2008).

Definition 44 *Let β be the angle with respect to the horizontal direction, i.e. the direction of the x-axis. This angle is considered and measured in radians $(0 \leq \beta \leq \pi)$. The values of β close to 0 and π are considered similar. Consider the distribution function*

$$\mathbb{P}_o(\beta \leq \varphi) \quad for \quad 0 \leq \varphi \leq \pi.$$

with corresponding probability density function $\vartheta(\varphi)$, satisfying

$$\mathbb{P}_o(\beta \leq \varphi) = \int\limits_0^{\varphi} \vartheta(\psi) \mathrm{d}\psi.$$

Then

$$\mathbb{P}_o(\varphi_1 \leq \beta \leq \varphi_2) = \int\limits_{\varphi_1}^{\varphi_2} \vartheta(\psi) \mathrm{d}\psi.$$

The quantity $\mathbb{P}_o(\varphi_1 \leq \beta \leq \varphi_2)$ is the proportion of the connecting lines with angles of orientation between r_1 and r_2. In the isotropic case $\vartheta(\varphi)$ is constant,

$$\vartheta(\varphi) \equiv \frac{1}{\pi}.$$

Second-order orientation analysis

The second-order orientation analysis is based on the reduced second-order moment measure $\mathcal{K}$ (Stoyan and Stoyan (1994)). These authors consider the sample of the orientations of all lines which connect pairs of points of an inter-point distance between r_1 and r_2 and determine the orientation distribution.

Definition 45 *The reduced second-order moment measure $\mathcal{K}$ in (1.4) can be described by the function $K(r, \alpha)$,*

$$K(r, \alpha) = \mathcal{K}(s(r, \alpha)),$$

where $s(r, \alpha)$ is the sector of radius r, centred at the origen and given by the angle α with respect to the x-axis.

For a fixed r, the ratio $K(r, \alpha)/K(r, \pi)$ is a distribution function, with a corresponding density function similar the above $\mathbb{P}_o(\beta \leq r)$.

For $r_1 > 0$, the description may be refined by introducing two distances r_1 and r_2 $(0 \leq r_1 < r_2)$ and considering the sector ring $s(r_1, r_2, \alpha)$ given by

$$s(r_1, r_2, \alpha) = \frac{s(r_2, \alpha)}{s(r_1, \alpha)}$$

The corresponding density function is denoted by $\vartheta_{(r_1, r_2)}(\varphi)$, which describes the distribution of the random orientations of the lines connecting the typical point and other points at distances between r_1 and r_2.

Lemma 46 *The point pair orientation distribution function can be expressed in terms of the reduced second-order moment measure by*

$$O_{(r_1, r_2)}(\varphi) = \frac{\mathcal{K}(s(r_1, r_2, \alpha))}{\mathcal{K}(b(\mathbf{0}, r_2)) - \mathcal{K}(b(\mathbf{0}, r_1))},$$

21

where it assumed that the denominator is positive. Then the following relation is satisfied

$$\vartheta_{(r_1,r_2)}(\varphi) = \int\limits_0^\varphi O_{(r_1,r_2)}(\psi)\mathrm{d}\psi,$$

for details see Stoyan and Stoyan (1994), Illian et al. (2008)).

1.5 Extension to multivariate processes

Let Y be a point process on $U \subseteq \mathbb{R}^d$. Given some space M, if a random *mark $m_\xi \in M$* is attached to each point $\xi \in Y$, then

$$X = \{(\xi, m_\xi) : \xi \in Y\}$$

is called a marked point process with points in U and mark space M. One simple example is a *multitype point process*, where $M = \{1, \ldots, n\}$ and the marks specify k different types of points. This is equivalent to a *k-dimensional multivariate point process*, that is a tuple $(X_1, \ldots, X_k)$ of point processes $X_1, \ldots, X_k$ corresponding to the k different types of points.

1.5.1 Marked Poisson processes

Consider a marked point process $X = \{(\xi, m_\xi) : \xi \in Y\}$ with points in U and mark space M.

Definition 47 *Suppose that Y is $Poisson(U, \phi)$, where ϕ is a locally integrable intensity function, and conditional on Y, the marks $\{m_\xi : \xi \in Y\}$ are mutually independent. Then X is a marked Poisson process. If the marks are identically distributed with a common distribution Q, then Q is called the mark distribution.*

Proposition 48 *Let X be a marked Poisson process with $M \in \mathcal{B}^p$ and where, conditional on Y, each mark m_ξ has a discrete or continuous density ρ_ξ which does not depend on $Y \backslash \{\xi\}$. Let $\rho(\xi, m) = \phi(\xi)\rho_\xi(m)$. Then*

i. $X \sim Poisson\,(U \times M, \rho)\,.$

ii. if the density on M defined by $\kappa\,(m) = \int \rho\,(\xi, m)\,\mathrm{d}\xi$ is locally integrable, then $\{m_\xi : \xi \in Y\} \sim Poisson\,(M, \kappa)$.

1.5.2 Multivariate Poisson processes

By a multivariate Poisson process it is usually meant that each X_i is a stationary Poisson process on $\mathbb{R}^d$ with intensity $0 < \rho_i < +\infty$ for $i = 1, \ldots, k$, see e.g. Diggle (2003, 2013). We have the equivalence between the following two properties:

1. $\mathbb{P}\,(m_\xi = i | Y = y) = p_\xi\,(i)$ depends only on ξ for realisations y of Y and $\xi \in y$.

2. $(X_1, \ldots, X_k)$ is a multivariate Poisson process with independent components $X_i \sim Poisson\,(U, \rho_i)$ where $\rho_i\,(\xi) = \phi\,(\xi)\,p_\xi\,(i)\,, i = 1, \ldots, k$.

A common hypothesis for marked point processes $\{m_\xi : \xi \in Y\}$ is that of random labelling which means that conditional on Y, the marks m_ξ are mutually independent and the distribution of m_ξ does not depend on Y. For a multitype Poisson process, for example, random labelling means 1. above with $\rho_\xi\,(i) = p\,(i)$ not depending on the location ξ.

1.5.3 Cross-moment measure

Consider a multivariate point process $X = (X_i)_{i=1}^k$, we assume that each X_i is a point process in $\mathbb{R}^d$ with intensity function ρ^i and count function N_i.

Definition 49 *Let $i, j \in \{1, \ldots, k\}$ be different, and set $a/0 = 0$, for $a \geq 0$.*

i. We define the cross-moment measure for points of types i and j by

$$\alpha^{ij}\,(C) = \mathbb{E}\left[\sum_{\xi \in X_i, \eta \in X_j} \mathbf{1}\,[(\xi, \eta) \in C]\right], \qquad C \in \mathcal{B}^d \times \mathcal{B}^d. \quad (1.22)$$

ii. If α^{ij} can be written as

$$\alpha^{ij}(C) = \int\limits_{\mathbb{R}^d} \int\limits_{\mathbb{R}^d} \mathbf{1}\left[(\xi, \eta) \in C\right] \rho^{ij(2)}(\xi, \eta) \, d\xi d\eta, \qquad C \in \mathcal{B}^d \times \mathcal{B}^d,$$

where $\rho^{ij(2)}$ is a non-negative function, then $\rho^{ij(2)}$ is called the cross second-order product density.

The cross pair correlation function for points of types i and j is defined by

$$g^{ij}(\xi, \eta) = \frac{\rho^{ij(2)}(\xi, \eta)}{\rho^i(\xi)\rho^j(\eta)}.$$

Definition 50 *Suppose that the measure*

$$\mathcal{K}^{ij}(B) = \frac{1}{|A|} \sum_{\xi \in X_i, \eta \in X_j} \frac{\mathbf{1}\left[\xi \in A, \eta - \xi \in B\right]}{\rho^i(\xi)\,\rho^j(\eta)}, \qquad B \in \mathcal{B}^d \qquad (1.23)$$

does not depend on the choice of $A \in \mathcal{B}^d$ with $0 < |A| < +\infty$. Then (X_i, X_j) is said to be cross second-order intensity reweighted stationary.

Consider a multivariate point process $X = (X_i)_{i=1}^k$, we assume that each X_i is a point process in $\mathbb{R}^d$ with intensity function ρ^i and count function N_i.

Definition 51 *The cross K- and L-functions are defined by*

$$K^{ij}(r) = \mathcal{K}^{ij}(b(\mathbf{0}, r)), \qquad L^{ij}(r) = \left(\frac{K^{ij}(r)}{\omega_d}\right)^{1/d} \qquad r > 0. \qquad (1.24)$$

Theorem 52 *If X_i and X_j are independent, then (X_i, X_j) is cross second-order*
reweighted stationary,

$$K^{ij}(B) = K^{ji}(B) = |B|, \qquad and \qquad L^{ij}(r) = r. \qquad (1.25)$$

Moreover,

$$K^{ij}(r, \alpha) = K^{ji}(r, \alpha).$$

Theorem 53 *Assume that (X_i, X_j) is stationary with intensities $0 < \rho^i <$ $+\infty$ and $0 < \rho^j < +\infty$. The nearest-neighbour function G^{ij} is the distribution function for the distance from a typical type i point to its nearest type j point, i.e.*

$$G^{ij}(r) = \frac{1}{\rho^i |A|} \mathbb{E} \left[\sum_{\xi \in X_i \cap A} \mathbf{1}[X_j \cap b(\xi, r) \neq \emptyset] \right], \qquad r > 0, \qquad (1.26)$$

for an arbitrary set $A \subset \mathbb{R}^d$ with $0 < |A| < +\infty$. Moreover, define

$$J^{ij}(r) = \frac{1 - G^{ij}(r)}{1 - F^j(r)}, \quad for \quad F^i(r) < 1.$$

The cross statistics G^{ij} and J^{ij} are not symmetric in i and j.

1.6   Non-parametric estimation

The non-parametric methods that do not assume a specific parametric model for the first-and second-order behaviour. The estimation approach is typically based on simple counts and kernel methods. We study methods for analyzing of spatial point pattern data are not linked to that a special families of parametric model.

In this section we let X be a spatial point process on $\mathbb{R}^d$ with intensity function ρ; if X is stationary, ρ is assumed to be a constant with $0 < \rho < +\infty$. Whenever needed we assume that the measure $\mathcal{K}$ exists. We confine to the case where a single point pattern $X_W = x$ is observed in a bounded window $W \in \mathcal{B}^d$ with $|W| > 0$, and discuss non-parametric estimation of ρ, K, g, L and J functions in the isotropic case. Also, we consider the anisotropic non-parametric estimate of the orientation distribution function. Higher-order summary statistics can be introduced as well, but the corresponding non-parametric estimators may be less stable if the number of points observed is not sufficiently large; see Peebles and Groth (1975), Stoyan and Stoyan (1994), Møller et al. (1998) and Schladitz and Baddeley (2000).

1.6.1 Estimation of intensity functions

In the homogeneous case, a natural unbiased estimate of the intensity function is

$$\hat{\rho} = \frac{N(X_W)}{|W|}. \tag{1.27}$$

This estimator is unbiased and, if N is ergodic, then it is consistent, i.e. as W increases it converges to the true value ρ. This holds independent of the specific distribution of N, whereas the variability of the estimator $\hat{\rho}$ is of course distribution-dependent (Illian et al. (2008)).

This is in fact the maximum likelihood estimate if X is a homogeneous Poisson process. In the inhomogeneous case, a non-parametric kernel estimate of the intensity function, following Diggle (1985), is

$$\hat{\rho}_\epsilon(\xi) = \sum_{\eta \in X_W} \frac{\kappa_\epsilon(\xi - \eta)}{c_{W,\epsilon}(\eta)}, \qquad \xi \in W. \tag{1.28}$$

Here κ_ϵ is a kernel with bandwidth $\epsilon > 0$, i.e.

$$\kappa_\epsilon(\xi) = \frac{1}{\epsilon^d} \kappa\left(\frac{\xi}{\epsilon}\right),$$

where κ is a given density function, and

$$c_{W,\epsilon}(\eta) = \int_W \kappa_\epsilon(\xi - \eta)\,\mathrm{d}\xi, \tag{1.29}$$

is an *edge-correction factor*.

The estimate (1.28) is usually sensitive to the choice of ϵ. It is usual when $d = 2$, to use a product kernel given by $\kappa(\xi) = e(\xi_1)e(\xi_2)$ for $\xi = (\xi_1, \xi_2) \in \mathbb{R}^d$, where

$$e(u) = \frac{3}{4}(1 - |u|)\,\mathbf{1}\left[|u| \leq 1\right], \qquad u \in \mathbb{R}, \tag{1.30}$$

is the *Epanechnikov kernel* (Epanechnikov (1969)).

Lemma 54 $\int_W \hat{\rho}_\epsilon(\xi)\,\mathrm{d}\xi$ *is an unbiased estimate of* $\mu(W)$.

A simple stationarity test is closely related to the intensity: determine statistically the intensity function $\rho(\cdot)$ and verify that a plot of the obtained estimate shows only local irregularities but not a general trend.

1.6.2 Estimation of K-, L- and g-functions

For non-parametric estimation of $\mathcal{K}$, it is useful to establish the following steps.

Notation 55 *Let* $|\Phi|_\alpha$ *denote the α-dimensional Hausdorff measure of* $\Phi \in \mathcal{B}^d$.
See Stoyan and Stoyan (1994).

Definition 56 *Let* $\xi, \eta \in W$, *we define the Ripley's isotropic edge-correction factor, (see Ripley (1976), Illian et al. (2008)) as*

$$w^d(\xi, \eta) = \frac{|\partial b(\xi, \|\xi - \eta\|) \cap W|_{d-1}}{|\partial b(\xi, \|\xi - \eta\|)|_{d-1}}. \tag{1.31}$$

Here $|\cdot|_0$ *and* $|\cdot|_1$ *are the zero- and one-dimensional Hausdorff measures respectively in* $\mathbb{R}$ *and* $\mathbb{R}^2$. *The zero-dimensional Hausdorff measure in* $\mathbb{R}$ *is the number of points of the set. The one-dimensional Hausdorff measure of a simple curve in* $\mathbb{R}^2$ *is equal to the length of the curve.*

Lemma 57 *Suppose that* X *is second-order intensity reweighted stationary. Then*

$$\sum_{\xi, \eta \in X_W}^{\neq} \frac{\mathbf{1}[\eta - \xi \in B]}{\rho(\xi)\rho(\eta)w^d(\xi, \eta)}, \tag{1.32}$$

is an unbiased estimator for $\mathcal{K}(B)$.

Lemma 57 provides an unbiased estimate of $\mathcal{K}(B)$ provided ρ is known. In practice ρ is not known, so $\rho(\xi)\rho(\eta)$ in (1.32) must be replaced by the estimator $\widehat{\rho(\xi)\rho(\eta)}$. The combined estimate

$$\hat{\mathcal{K}}(B) = \sum_{\xi, \eta \in X_W}^{\neq} \frac{\mathbf{1}[\eta - \xi \in B]}{\widehat{\rho(\xi)\rho(\eta)}w^d(\xi, \eta)}, \tag{1.33}$$

is then biased (see Illian et al. (2008)).

In fact unbiasedness is usually unobtainable for many estimators in spatial statistics, but instead they are often ratio-unbiased, i.e. of the form $\hat{\theta} = Y/Z$ where $\theta = \mathbb{E}[Y]/\mathbb{E}[Z]$. For example, in the homogeneous case, if $\widehat{\rho(\xi)\rho(\eta)} = \widehat{\rho^2}$ is unbiased, then (1.33) is ratio-unbiased. Stoyan and Stoyan (2000) discuss various possibilities for the homogeneous case: one possibility is to transform the estimate in (1.27) to obtain $\left(N(X_W)\right)^2/|W|^2$ as an estimate of ρ^2; an alternative is

$$\widehat{\rho^2} = \frac{N(X_W)(N(X_W)-1)}{|W|^2}, \qquad (1.34)$$

which is unbiased for a Poisson process. For the inhomogeneous case, Baddeley et al. (2000) propose to use $\widehat{\rho(\xi)\rho(\eta)} = \bar{\rho}_\epsilon(\xi)\bar{\rho}_\epsilon(\eta)$ where

$$\bar{\rho}_\epsilon(\xi) = \sum_{\eta \in X_W \setminus \{\xi\}} \frac{\kappa_\epsilon(\xi - \eta)}{c_{W,\epsilon}(\eta)}, \qquad \xi \in W, \qquad (1.35)$$

is a slight modification of (1.28). Also, Baddeley et al. (2000) show that for an inhomogeneous Poisson processes, $\bar{\rho}_\epsilon(\xi)$ is less biased than $\hat{\rho}_\epsilon(\xi)$ when $\xi \in X_W$ is a data point.

Remark 58 *Consider* $\hat{K}(r) = \hat{\mathcal{K}}(b(\mathbf{0}, r))$:

1. *From their definition, it is clear that* $1/w^d(\xi, \eta) \geq 1$. *Also, the larger weights tend to be associated with pairs of events separated by large distances. Typically,* $\mathbb{V}ar\{\hat{K}(r)\}$ *tends to increase with* r.

2. *The dimensions of* W *clearly limit the range of values of* r *which can be considered. In practice, the increasing variance of* $\hat{K}(r)$ *is a more serious limitation. As a rough guide, for data on a rectangle* W, *it is usually not worth trying to estimate* $K(r)$ *at values of* r *bigger than one-half the length of the shorter side of* W.

The estimate of $L(r)$ obtained from transforming that of $K(r)$ is in general biased.

Lemma 59 *An alternative edge-correction kernes estimator for the pair correlation function (Fiksel (1988a), Stoyan and Stoyan (1994), Baddeley et al. (2000)) is given by*

$$\hat{g}(r) = \frac{1}{\sigma_d r^d |W|} \sum_{\xi,\eta \in X_W}^{\neq} \frac{\kappa_b(\|\xi - \eta\| - r)}{\widehat{\rho(\xi)\rho(\eta)} w^d(\xi,\eta)}. \tag{1.36}$$

Here $\kappa_b(u) = \kappa(u/b)/b$, $u \in \mathbb{R}$ and bandwidth $b > 0$, (see Illian et al. (2008)).

1.6.3 Estimation of F-, G- and J-functions

Reduced-sample estimators of F and G given in (1.20) and (1.21) are derived using minus sampling.

Definition 60 *Let*

$$d(\xi, B) = \inf\{\|\xi - \eta\| \, \eta \in B\},$$

be the shortest distance from a point $\xi \in \mathbb{R}^d$ to a set $B \subset \mathbb{R}^d$. Let $I \subset \mathbb{R}^d$ denote a finite regular grid of points (chosen independently of X), and let $\#I_r$ denote the cardinality of the set $I_r = I \cap W_\ominus r$, where $W_\ominus r = \{\xi \in W : b(\xi, r) \subseteq W\}$ for $r > 0$.

Lemma 61 *The following estimate is unbiased for F*

$$\hat{F}(r) = \sum_{\xi \in I_r} \frac{\mathbf{1}[d(\xi, X_W) \leq r]}{\#I_r},$$

for $\#I_r > 0$, and the next one is a ratio-unbiased estimate for G

$$\hat{G}(r) = \sum_{\xi \in X_W \cap W_\ominus} \frac{\mathbf{1}[d(\xi, X_W) \leq r]}{\hat{\rho}|W_\ominus|},$$

for $|W_\ominus| > 0$. Finding

$$\hat{J}(r) = \frac{1 - \hat{G}(r)}{1 - \hat{F}(r)}, \quad for \quad \hat{F}(r) < 1.$$

1.6.4 Estimation for directional statistics

For point patterns with a small number of points this procedure is not possible. In the planar case, the estimator for the orientation distribution function is given by

$$\hat{O}_{(r_1,r_2)}(\varphi) = \frac{\hat{K}(s(r_1, r_2, \varphi))}{\hat{K}(r_2) - \hat{K}(r_1)}, \qquad 0 \le \varphi \le \pi.$$

The corresponding edge-corrected unbiased estimator of the point pair oriantation function $\vartheta_{(r_1,r_2)}(\varphi)$ is given by

$$\hat{\vartheta}_{(r_1,r_2)}(\varphi) = \sum_{\xi,\eta \in X_W}^{\neq} \frac{\mathbf{1}[r_1 \le \|\xi - \eta\| \le r_2]\kappa(\varphi - \beta_{\xi\eta})}{w^2(\xi,\eta)}, \qquad 0 \le \varphi \le \pi,$$

where κ is the Epanechnikov kernel function and $\beta_{\xi\eta}$ the orientation angle of the line through the points ξ and η (see Illian et al. (2008)).

1.6.5 Multivariate estimators

Lemma 62 *If (X_i, X_j) is cross second-order reweighted stationary, then for $B \in \mathcal{B}^d$*

$$\sum_{\substack{\xi \in X_i \cap W \\ \eta \in X_j \cap W}}^{\neq} \frac{\mathbf{1}[\eta - \xi \in B]}{\rho_i(\xi)\rho_j(\eta)w^d(\xi,\eta)}, \tag{1.37}$$

is an unbiased estimator of $\mathcal{K}^{ij}(B)$.

In (1.37) we can substitute $\rho_i(\xi)\rho_j(\eta)$ with a non-parametric estimator $\widehat{\rho_i(\xi)\rho_j}(\eta)$.

By **Theorem 52** if $X = (X_1, \ldots, X_k)$ is a multivariate point process, then the $k(k+1)/2$ functions $K^{ij}(r)$ with $1 \le i \le j < k$, completely describe the *second-order* properties of the process. A necessary, but not

30

sufficient, condition for the process of points of type i to be independent of the process of points of type j is that $K^{ij}(r) = \omega_d r^d$ for all r. Not surprisingly, description and estimation of the second-order structure of a multitype process requires consideration of only two of the types at a time. Therefore it will be sufficient when discussing the problem of estimation to consider a two-type process consisting of n_1 points $\{\varkappa_1, \ldots, \varkappa_{n_1}\}$ and n_2 points $\{\varsigma_1, \ldots, \varsigma_{n_2}\}$. Suppose that we observe such a process over a plane region W.

Lotwick and Silverman (1982), and Diggle (2003), give an combined estimator for the bivariate K- function:

$$\hat{K}^{ij}(r) = \frac{|W|}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \frac{\mathbf{1}[\|\varkappa_i - \varkappa_j\| \le r]}{w^{d*}(\varkappa_i, \varsigma_j)}, \qquad (1.38)$$

where

$$w^{d*}(\varkappa_i, \varsigma_j) = \frac{n_1 w^d(\varsigma_i, \varkappa_j) + n_2 w^d(\varkappa_i, \varsigma_j)}{n_1 + n_2}.$$

If g^{ij} is isotropic, a kernel estimator for the cross pair correlation is given by

$$\hat{g}^{ij}(r) = \frac{1}{\sigma_d r^d |W|} \sum_{\xi, \eta \in X_W}^{\ne} \frac{\kappa_b(\|\xi - \eta\| - r)}{\widehat{\rho^i(\xi)} \, \widehat{\rho^j}(\eta) w^d(\xi, \eta)}.$$

Similarly, for (X_i, X_j) the reduced-sample estimator of $G^{ij}(r)$ is

$$\hat{G}^{ij}(r) = \sum_{\xi \in X_i \cap W_{\ominus r}} \frac{\mathbf{1}[d(\xi, X_j \cap W) \le r]}{\hat{\rho}^i |W_{\ominus}|},$$

By substitution of $\widehat{F}^j(r)$ and $\widehat{G}^{ij}(r)$ functions the estimator for $\hat{J}^{ij}(r)$ is obtained.

1.7   Envelopes procedure

We describe this procedure following Møller and Waagepetersen (2004). Consider a simple hypothesis H_0. Confidence intervals and other distributional characteristics associated with the non-parametric estimate $\hat{R}$ can

be obtained by a *bootstrap* using simulation under H_0. For a given distance $\xi \in B$, let $T_0(\xi) = T(X, \xi)$ denote any statistic obtained from the point process X observed within the window W. Let $\Theta = \{T_i(\xi)\}_{i=1}^n$ be obtained from *i.i.d.* simulations $X_1, \ldots, X_n$ under H_0. From the empirical distribution of Θ we can estimate any quantile for the distribution of $T_0(\xi)$ under H_0, and we can do this with any desired precision if n is large enough. Notice that although T_i and T_j are *i.i.d.*, the random vectors $(T_1(\xi), \ldots, T_n(\xi))$ considered for different values of ξ are dependent. So some caution should be taken when we compare the results for different values of ξ.

If the computation of $T_i(\xi)$, $i = 1, \ldots, n$, is time consuming, the following envelopes may be used where n is small. Let

$$T_{\min}(\xi) = \min \Theta, \qquad \text{and} \qquad T_{\max}(\xi) = \max \Theta, \qquad (1.39)$$

under H_0,

$$\mathbb{P}(T_0(\xi) < T_{\min}(\xi)) = \mathbb{P}(T_0(\xi) > T_{\max}(\xi)) \le \frac{1}{n+1}, \qquad (1.40)$$

with equality if $T_0(\xi), T_1(\xi), \ldots, T_n(\xi)$ are almost surely different. The bounds $T_{\min}(\xi)$ and $T_{\max}(\xi)$ are called the $100/(n+1)\%$-lower and the $100n/(n+1)\%$-upper envelopes.

1.8 Spatio-temporal point processes

Spatio-temporal point processes are considered as a hybrid of spatial and temporal components by extending the definition of spatial point process to include time.

Spatio-temporal point processes have been studied thoroughly in the context of earthquake data by Ogata (1998), who wrote a summary paper of parametric, maximum likelihood techniques. Choi and Hall (2001) added non-parametric estimators of the intensity function using a kernel estimator approach, and discuss asymptotic theory for many parametric

estimators. Rathbun and Cressie (1994) discuss spatio-temporal point processes in the context of tree growth.

Because the spatial location can always be considered as one component of a multi-dimensional mark (see Daley and Vere-Jones (2008)), the evolution of spatial features with time is often of special interest. Despite such considerations, studies of spatio-temporal models have lagged well behind those of simple temporal models, and even those of purely spatial models. No doubt the reasons have been largely practical, notably the difficulty of compiling good spatio-temporal datasets and the heavy computations needed to analyse them.

One way to observe those processes developed by Daley and Vere-Jones (2008) is to consider the spatial location itself viewed as a mark for a simple point process in time, thereby providing one route to likelihood analyses of spatio-temporal models. Further characteristics, such as magnitude, spatial extent, or even duration, can be added as additional marks. Thus, the study of spatio-temporal point processes leads almost inevitably to the more general study of evolving spatial fields, although practical modelling in this direction is still limited and very subject-specific.

Following our approach, Diggle et al. (1995) consider the problem of detecting and describing spatio-temporal interactions in point process data. They extend existing second-order methods for purely spatial point process data to the spatial-temporal setting. This extension allows to estimate spatio-temporal interaction as a function of spatial and temporal separation. Gabriel and Diggle (2009) extend to the spatio-temporal setting a method proposed by Baddeley et al. (2000) for inhomogeneous spatial point process data. Møller and Ghorbani (2012) consider second-order analysis based on pair correlation functions and K-functions for general inhomogeneous spatio-temporal point processes assuming spatio-temporal separability of the intensity function, but clarify different meanings of second-order spatio-temporal separability. One is second-order spatio-temporal independence, and another concerns a separable spatio-temporal covariance density. Ghorbani (2013) suggests a weak stationarity of a spatio-temporal point process test and Gabriel (2013) builds a rigorous simulation study to

show the efficiency of the second-order estimators on differents scenarios for various spatio-temporal edge-corrections.

Cronie and Lieshout (2014) extend and estimate the J-function for inhomogeneous spatio-temporal point processes. Tamayo-Uria et al. (2014) analyse the spatio-temporal distribution of rat sightings, which are directly related to rat infestation. They formulate a more mechanistic model in which the conditional intensity of the point process depends explicitly on its past history. Diggle (2013) is the recent book on this methodology with nice examples in epidemiology.

1.8.1 First- and second-order spatio-temporal measures

Definition 63 *Let X be a random countable subset of $\mathbb{R}^2 \times \mathbb{R}$. Consider $W \subset \mathbb{R}^2$ a bounded spatial region $|W| > 0$ and a bounded time interval $|T| > 0$, and $X \cap (W \times T) = \{(\mathbf{u}_i, s_i), i = 1, \ldots, n\}$ stands for data. Let N_{lfs} and N_{lft} be the spaces of locally finite subsets of $\mathbb{R}^2$ and $\mathbb{R}$ equipped with σ-algebras $\mathcal{N}_{\mathrm{lfs}}$ and $\mathcal{N}_{\mathrm{lft}}$ respectively. In the sequel, $N(A)$ denotes the number of the events of the process falling in a bounded region $A \subset W \times T$.*

For a rigurous definition of a point process based on measure theory see e.g. Daley and Vere-Jones (2008).

Definition 64 *For a given event $(\mathbf{u}, s)$, the events that are* close *to $(\mathbf{u}, s)$ in both space and time, for each spatial distance r, and time lag t, are given by the corresponding spatio-temporal cylindrical neighborhood of the event $(\mathbf{u}, s)$, which can be expressed by the cartesian product as*

$$b((\mathbf{u}, s), r, t) = \{(\mathbf{v}, l) : \|\mathbf{u} - \mathbf{v}\| \le r, |s - l| \le t\}, \qquad (\mathbf{u}, s), (\mathbf{v}, l) \in \mathbb{R}^2 \times \mathbb{R},$$

where $\| \cdot \|$ denotes the Euclidean distance in $\mathbb{R}^2$ and $|\cdot|$ denotes the usual distance in $\mathbb{R}$. Note that $b((\mathbf{u}, s), r, t)$ is a cylinder with center $(\mathbf{u}, s)$, radius r and height $2t$.

Theorem 65 *Assume that X has spatio-temporal intensity function ρ and spatio-temporal pair correlation function g. Then by (1.1)*

$$\int h(\mathbf{u}, s)\rho(\mathbf{u}, s)\mathrm{d}(\mathbf{u}, s) = \mathbb{E} \sum_{(\mathbf{u},s)\in X} h(\mathbf{u}, s),$$

and

$$\iint f((\mathbf{u}, s), (\mathbf{v}, l))g((\mathbf{u}, s), (\mathbf{v}, l))\mathrm{d}(\mathbf{u}, s)\mathrm{d}(\mathbf{v}, l) =$$

$$\mathbb{E} \sum_{(\mathbf{u},s),(\mathbf{v},l)\in X}^{\neq} \frac{f((\mathbf{u}, s), (\mathbf{v}, l))}{\rho(\mathbf{u}, s)\rho(\mathbf{v}, l)}, \quad (1.41)$$

for any non-negative Borel functions h and f defined on $\mathbb{R}^2 \times \mathbb{R}$ and $(\mathbb{R}^2 \times \mathbb{R}) \times (\mathbb{R}^2 \times \mathbb{R})$ respectively. Here $\sum^{\neq}$ means that $(\mathbf{u}, s) \neq (\mathbf{v}, l)$, and we take $a/0 = 0$ for $a \geq 0$.

It follows from Equation (1.41) that with probability one, for any pair of distinct points $(\mathbf{u}, s)$ and $(\mathbf{v}, t)$ from X, we have that $\mathbf{u} \neq \mathbf{v}$ and $s \neq l$ (see Møller and Ghorbani (2012)). Then we can ignore the case where the spatial and temporal component processes (say X_{space} and X_{time}) have multiple points, and define them by

$$X_{\mathrm{space}} = \{\mathbf{u} : (\mathbf{u}, s) \in X, s \in T\}, \qquad X_{\mathrm{time}} = \{s : (\mathbf{u}, s) \in X, \mathbf{u} \in W\}.$$

We consider X_{space} and X_{time} rather than the marginal processes given by all events respective all times in X, since the later processes could not have well-defined first- and second-order properties (see Møller and Ghorbani (2012)).

1.8.2 First-order spatio-temporal separability

We make the pragmatic working assumption that first-order effects are separable (see Møller and Ghorbani (2012)), meaning that if $(\mathbf{u}, s) \in \mathbb{R}^2 \times \mathbb{R}$, $\rho(\mathbf{u}, s)$ can be factorised as

$$\rho\left(\mathbf{u}, s\right) = \bar{\rho}_1\left(\mathbf{u}\right)\bar{\rho}\left(s\right), \tag{1.42}$$

where $\bar{\rho}_1$ and $\bar{\rho}_2$ are non-negative functions. Under this assumption, any non-separable effects are interpreted as second-order, rather than first-order. We have

$$\begin{aligned} \mu\left(A \times B\right) &= \mathbb{E}\left[N\left(A \times B\right)\right] \\ &= \int_A \bar{\rho}_1\left(\mathbf{u}\right)\mathrm{d}\mathbf{u} \int_B \bar{\rho}_2\left(s\right)\mathrm{d}s. \end{aligned}$$

Definition 66 *Define the marginal spatial and temporal intensity functions ρ_{space} and ρ_{time} respectively as*

$$\rho_{space}\left(\mathbf{u}\right) = \bar{\rho}_1\left(\mathbf{u}\right)\int_T \bar{\rho}_2\left(s\right)\mathrm{d}s, \qquad \rho_{time}\left(s\right) = \bar{\rho}_2\left(s\right)\int_W \bar{\rho}_1\left(\mathbf{u}\right)d\mathbf{u}.$$

Remark 67 *Note that ρ satisfies:*

$$\rho\left(\mathbf{u}, s\right) \propto \rho_{space}\left(\mathbf{u}\right)\rho_{time}\left(s\right).$$

If X is stationary, ρ, ρ_{space} and ρ_{time} are all constant.

1.8.3 Palm distribution

Assume we have spatio-temporal point process $X \subseteq \mathbb{R}^2 \times \mathbb{R}$ with intensity measure μ (see Møller and Ghorbani (2012)).

Definition 68 *The Campbell measure is defined by the relation*

$$C(B \times F) = \mathbb{E}\left[\mathbf{1}[X \in F]\right]N(B), \qquad B \subseteq \mathbb{R}^2 \times \mathbb{R}, \qquad F \in \mathcal{N}_{\mathrm{lfs}} \times \mathcal{N}_{\mathrm{lft}},$$

where $B = W \times T$.

Note that $\mu(\cdot) = C(\cdot \times (N_{\mathrm{lfs}} \times N_{\mathrm{lft}}))$. For each $B \subseteq \mathbb{R}^2 \times \mathbb{R}$ and $F \in \mathcal{N}_{\mathrm{lfs}} \times \mathcal{N}_{\mathrm{lft}}$, $\mu(B) = 0$ then $C(B \times F) = 0$ therefore $C(\cdot \times F)$ is absolutely

continuous wiht respect to μ. Then under the Radon-Nikodym theorem, there exists a μ-almost surely unique density $(\mathbf{u}, s) \to P_{(\mathbf{u}, s)}(F)$ so that

$$C(B \times F) = \int_B P_{(\mathbf{u}, s)}(F)\mu(\mathrm{d}(\mathbf{u}, s)), \qquad (1.43)$$

where $P_{(\mathbf{u}, s)}$ is the Palm distribution. In the case of stationary point processes the collection of Palm distributions is determined by P_o in the origin $P_{(\mathbf{u}, s)}(\cdot) = P_o(\cdot - (\mathbf{u}, s))$ for $(\mathbf{u}, s) \in \mathbb{R}^2 \times \mathbb{R}$.

Definition 69 *The reduced Campbell measure $C^!$ on $(\mathbb{R}^2 \times \mathbb{R}) \times (\mathcal{N}_{\mathrm{lfs}} \times \mathcal{N}_{\mathrm{lft}})$ is defined by*

$$C^!(D) = \mathbb{E}\left[\sum_{(\mathbf{u}, s) \in X} \mathbf{1}[((\mathbf{u}, s), X \setminus \{(\mathbf{u}, s)\}) \in D] \right].$$

Analogously we can define the distribution $P^!_{(\mathbf{u}, s)}$ called reduced Palm distribution and it can be interpreted as the conditional distribution of a point process given that $(\mathbf{u}, s)$ is a point of the process. Further, it satisfies the same relation in (1.43). For a Poisson point process, the reduced Palm distribution is the same as the original distribution of the process.

Theorem 70 *For a spatio-temporal point process X and any non-negative Borel function h,*

$$\mathbb{E}\left[\sum_{(\mathbf{u}, s) \in X} h((\mathbf{u}, s), X \setminus \{(\mathbf{u}, s)\}) \right]$$

$$= \int_{N_{\mathrm{lfs}} \times N_{\mathrm{lft}}} \int_{\mathbb{R}^2 \times \mathbb{R}} h((\mathbf{u}, s), (x, \zeta))C^!(\mathrm{d}(\mathbf{u}, s), \mathrm{d}(x, \zeta))$$

$$= \int_{\mathbb{R}^2 \times \mathbb{R}} \int_{N_{\mathrm{lfs}} \times N_{\mathrm{lft}}} h((\mathbf{u}, s), (x, \zeta))P^!_{(\mathbf{u}, s)}(\mathrm{d}(x, \zeta))\, \mu(\mathrm{d}(\mathbf{u}, s)),$$

where $(x, \zeta) \in N_{\mathrm{lfs}} \times N_{\mathrm{lft}}$.

Theorem 71 *For a stationary spatio-temporal point process X with intensity ρ and for any non-negative Borel function h*

$$\mathbb{E}\left[\sum_{(\mathbf{u},s)\in X} h((\mathbf{u},s), X \setminus \{(\mathbf{u},s)\})\right]$$
$$= \rho \int_{\mathbb{R}^2\times\mathbb{R}} \int_{N_{\mathrm{lfs}}\times N_{\mathrm{lft}}} h((\mathbf{u},s),(\mathbf{u}+x,s+\zeta))P_o^!(\mathrm{d}(x,\zeta))\,\mathrm{d}(\mathbf{u},s),$$

where $(\mathbf{u}+x,s+\zeta) = \{(z,\eta) + (\mathbf{u},s) : z \in x \text{ and } \eta \in \zeta\}$ denote the translation of the point configuration (x,ζ) by $(\mathbf{u},s)$.

For more details see Daley and Vere-Jones (2003), Møller and Waagepetersen (2004) and Chiu et al. (2013).

1.8.4 Second-order characteristics

Following the definition of the K-function due to Møller and Ghorbani (2012) according to the spatio-temporal context as above introduced, we have the following definition and proposition

Definition 72 *Let X be a second-order intensity-reweighted stationarity SOIRS in the sense established by Baddeley et al. (2000), then*

$$K(r,t) = \int \mathbf{1}\left[\|\mathbf{u}\| \leq r, |s| \leq t\right] g(\mathbf{u},s)\,d(\mathbf{u},s), \qquad r > 0, \quad t > 0. \quad (1.44)$$

In the stationary case of X, $\rho K(r,t)$ is the expected number of further points within distance r and time lag t from the origin given that X has a point at the origin (see Ripley (1976) and Ripley (1977)). There is another weak definition in a more heuristic way provided by Gabriel and Diggle (2009) and which takes only the present and future events, but its value only differs by a factor of $1/2$.

Definition 73 *For a SOIRS, isotropic and stationary spatio-temporal point process X, the spatio-temporal pair correlation function is proportional to*

the derivative of $K(r,t)$ with respect to r and t, then

$$g(r,t) = \frac{1}{4\pi r}\frac{\partial^2 K(r,t)}{\partial r \partial t}, \qquad r > 0, \quad t > 0. \qquad (1.45)$$

Proposition 74 *If X is a Poisson process,*

$$K(r,t) = 2\pi r^2 t, \qquad and \qquad g(r,t) = 1.$$

When a spatio-temporal point process holds the above condition, named as *complete spatio-temporal randomness.*

1.8.5   Second-order spatio-temporal separability

Definition 75 *The hypothesis of spatio-temporal separability of the pair correlation function states that $g((\mathbf{u}, s), (\mathbf{v}, l)) = \bar{g}_1(\mathbf{u}, \mathbf{v})\bar{g}_2(s, l)$ where $\bar{g}_1$ and $\bar{g}_2$ are non-negative functions.*

Intuitively, the probability of observing a pair of points from X occurring jointly in each of two infinitesimally small sets with centers $(\mathbf{u}, s), (\mathbf{v}, l)$ and volumes $d(\mathbf{u}, s)$, $d(\mathbf{v}, l)$ is

$$[\bar{\rho}_1(\mathbf{u})\bar{\rho}_1(\mathbf{v})\bar{g}_1(\mathbf{u}, \mathbf{v})\mathrm{d}\mathbf{u}\mathrm{d}\mathbf{v}][\bar{\rho}_2(s)\bar{\rho}_2(l)\bar{g}_2(s, l)\mathrm{d}s\mathrm{d}l]$$

which is a product of a function of the locations $(\mathbf{u}, \mathbf{v})$ and the areas $(\mathrm{d}\mathbf{u}, \mathrm{d}\mathbf{v})$ and a function depending on the times (s, l) and the lengths $(\mathrm{d}s, \mathrm{d}l)$.

Following Møller and Ghorbani (2012), we can write the spatial and temporal components of the K-function as

$$K_{\text{space}}(r) = \int_{\|\mathbf{u}\|\le r} g_{\text{space}}(\mathbf{u})\,\mathrm{d}\mathbf{u} \ \text{ and } \ K_{\text{time}}(t) = \int_{-t}^{t} g_{\text{time}}(s)\,\mathrm{d}s, \ r > 0, t > 0,$$

such that,

$$g_{\text{space}}(\mathbf{u}, \mathbf{v}) = g_{\text{space}}(\mathbf{u} - \mathbf{v}) = \int_T \int_T p_1(s) p_2(l) g(\mathbf{u} - \mathbf{v}, s - l) \mathrm{d}s \mathrm{d}l$$

and

$$g_{\text{time}}(s, l) = g_{\text{time}}(s - l) = \int_W \int_W p_1(\mathbf{u}) p_2(\mathbf{v}) g(\mathbf{u} - \mathbf{v}, s - l) \mathrm{d}\mathbf{u} \mathrm{d}\mathbf{v},$$

with

$$p_1(\mathbf{u}) = \frac{\bar{\rho}_1(\mathbf{u})}{\int_W \bar{\rho}_1(\mathbf{u}) \mathrm{d}\mathbf{u}}, \quad \text{and} \quad p_2(s) = \frac{\bar{\rho}_2(s)}{\int_T \bar{\rho}_2(l) \mathrm{d}l}.$$

1.8.6 Spatio-temporal inhomogeneus J-function

Cronie and Lieshout (2014) recently submitted the first extension of inhomogeneous J-function in the context of intensity-reweighted moment stationary spatio-temporal point processes. They build their statistic in terms of the supremum metric which is topologically equivalent to the Euclidean space. We follow them in the sequel

Definition 76 *Let X be a spatio-temporal point process for which the second-order product density exists. If $\bar{\rho} = \inf_{(\mathbf{u}, s)} \rho(\mathbf{u}, s) > 0$ and for $n \geq 1$, ξ_n is translation invariant in the sense that*

$$\xi_n((\mathbf{u}_1, s_1) + (a, b), \ldots, (\mathbf{u}_n, s_n) + (a, b)) = \xi_n((\mathbf{u}_1, s_1), \ldots, (\mathbf{u}_n, s_n)),$$

for almost all $(\mathbf{u}_1, s_1), \ldots, (\mathbf{u}_n, s_n) \in \mathbb{R}^2 \times \mathbb{R}$ and all $(a, b) \in \mathbb{R}^2 \times \mathbb{R}$, we say that X is intensity-reweighted moment stationary (IRMS).

Definition 77 *The generating function $G(\cdot)$ on X is defined as*

$$G(v(\mathbf{u}, s)) = \mathbb{E}\left[\prod_{(\mathbf{u}, s) \in X} v(\mathbf{u}, s) \right] = \int \prod_{(\mathbf{u}, s) \in X} v(\mathbf{u}, s) P_o^!(\mathrm{d}(x, \zeta)) \, \mathrm{d}(\mathbf{u}, s),$$

for all functions $v = 1 - u$ such that $u : \mathbb{R}^2 \times \mathbb{R} \to [0, 1]$ is measurable with bounded support on $\mathbb{R}^2 \times \mathbb{R}$.

By convention, an empty product equals 1. The generating functional uniquely determines the distribution of X.

The stationary J-function

Definition 78 *Assume that X is stationary, so we define analogously as in the spatial case*

$$J(r,t) = \frac{1 - G(r,t)}{1 - F(r,t)} = \frac{\mathbb{P}_0^!(X \cap S_{st} \neq \emptyset)}{\mathbb{P}(X \cap S_{st} \neq \emptyset)}$$

for $r \geq 0$, $t \geq 0$ such that $F(r,t) \neq 1$, where

$$S_{st} = \{(\mathbf{u}, s) : \|\mathbf{u}\| \leq r, |s| \leq t\}, \qquad (\mathbf{u}, s) \in \mathbb{R}^2 \times \mathbb{R},$$

with $\mathbb{P}_0^!$ reversely induced probability measure $\mathcal{F}$.

The inhomogeneous J-function

Definition 79 *Let X be an IRMS spatio-temporal point process. For $r \geq 0$, $t \geq 0$, let*

$$J_n(r,t) = \int_{S_{st}} \cdots \int_{S_{st}} \xi_{n+1}((\mathbf{0},0),(\mathbf{u}_1,s_1),\ldots,(\mathbf{u}_n,s_n)) \prod_{i=1}^{n} \mathrm{d}(\mathbf{u}_i,s_i)$$

and set

$$J_{inhom}(r,t) = 1 + \sum_{n=1}^{\infty} \frac{(-\bar{\rho})^n}{n!} J_n(r,t)$$

for all spatial ranges $r \geq 0$ and temporal ranges $t \geq 0$ for which the series is absolutely convergent.

1.8.7 Spatio-temporal models

Spatio-temporal models are a natural extension of spatial models considered in Section 1.3. To set the suitable context we focus on the definitions given by Gabriel et al. (2012).

Homogeneous spatio-temporal Poisson process

The homogeneous Poisson process is the simplest possible stochastic mechanism for the generation of spatio-temporal point patterns. It is rarely plausible as a model for data, but provides a benchmark of complete spatio-temporal randomness (CSTR). Informally, in a realisation of a homogenous Poisson process on any spatio-temporal region $W \times T$, the events form an independent random sample from the uniform distribution on $W \times T$. More formally, the homogeneous Poisson process is defined by the following postulates:

i. For some $\rho > 0$, the number $N(W \times T)$ of events within the region $W \times T$ follows a Poisson distribution with mean $\rho |W| |T|$ where $|\cdot|$ denotes (two-dimensional or one-dimensional) Lebesgue measure according to the context.

ii. Given $N(W \times T) = n$, the n events in $W \times T$ form an independent random sample from the uniform distribution on $W \times T$.

The first-order and second-order intensities of a homogeneous Poisson process reduce to constants, $(\mathbf{u}, s), (\mathbf{v}, l) \in W \times T$, then $\rho(\mathbf{u}, s) = \rho$ and $\rho^{(2)}((\mathbf{u}, s), (\mathbf{v}, l)) = \rho^2$.

Spatio-temporal Cox process

The Cox process or inhomogeneous Poisson process is the simplest non-stationary point process. It is obtained replacing the constant intensity of a homogeneous Poisson process by a spatially and/or temporally varying intensity function $\rho(\mathbf{u}, s)$, for $(\mathbf{u}, s) \in W \times T$. Inhomogeneous Poisson processes are defined by the following postulates:

i. The number $N(W \times T)$ of events within the region $W \times T$ follows a Poisson distribution with mean

$$\int_{W \times T} \rho(\mathbf{u}, s) \mathrm{d}(\mathbf{u}, s).$$

ii. Given $N(W \times T) = n$, the n events in $W \times T$ form an independent random sample from the distribution on $W \times T$ with probability density function

$$f(\mathbf{u}, s) = \frac{\rho(\mathbf{u}, s)}{\int_{W \times T} \rho(\mathbf{v}, s)\mathrm{d}(\mathbf{v}, s)}.$$

Spatio-temporal Neyman-Scott process

We define a spatio-temporal Poisson cluster process as the following direct generalisation of its spatial counterpart (Neyman and Scott (1958)).

i. Parents form a Poisson process with intensity $\rho_{\mathrm{p}}(\mathbf{u}, s)$.

ii. The number of offspring per parent is a random variable N_c with mean m_c, realised independently for each parent.

iii. The locations and times of the offspring relative to their parents are independently and identically distributed according to a trivariate probability density function $g : \mathbb{R}^2 \times \mathbb{R}_+ \to \mathbb{R}$.

iv. The final process is composed of the superposition of the offspringd only.

Spatio-temporal stationary Poisson cluster processes

The spatial distribution of the offsprings is a zero-mean bivariate isotropic normal distribution with standard deviation σ and the temporal distribution is exponential with rate α. The expected number of offsprings per parent follows a Poisson distribution with mean m_c. This process is an interpretation of a spatio-temporal shot-noise Cox process (see Møller and Diaz-Avalos (2010), Møller (2003), Gabriel (2013)) with residual process

$$R(\mathbf{u}, s) = \frac{1}{\rho} \sum_{(\mathbf{v}, l) \in X} \varphi(\mathbf{u} - \mathbf{v}, s - l),$$

where X is a stationary Poisson process in $\mathbb{R}^2 \times \mathbb{R}$ with intensity ρ and φ is the density function

$$\varphi(u,v) = \phi_{\sigma^2}^{(2)}(\|\mathbf{u} - \mathbf{v}\|)\mathcal{E}_\alpha(|s - l|),$$

here, $\| \cdot \|$ denotes the usual distance in $\mathbb{R}^2$ and $|\cdot|$ is the absolute value in $\mathbb{R}$, with the abuse of the notations u and v for $u = \|\mathbf{u} - \mathbf{v}\|$ and $v = |s - l|$. For such a process, we have $g(u,v) = 1 + \frac{1}{\rho}\varphi * \widetilde{\varphi}(u,v)$, where $*$ denotes convolution and $\widetilde{\varphi}(u,v) = \varphi(-u,-v)$. Then,

$$g(u,v) = 1 + \frac{\alpha}{8\pi\sigma^2\rho}\exp\left(-\frac{u^2}{4\sigma^2} - \alpha v\right),$$

and

$$K(u,v) = 2\pi u^2 v + \frac{1}{2\rho}(\exp(\alpha v) - \exp(-\alpha v))\left(1 - \exp\left(-\frac{u^2}{4\sigma}\right)\right).$$

Spatio-temporal geometric anisotropic Poisson cluster processes

This is defined as spatio-temporal stationary Poisson cluster processes but with

$$g(\mathbf{u}, s) = g_0(\sqrt{\mathbf{u}\Sigma^{-1}\mathbf{u}^t}, s),$$

where $\mathbf{u} \in \mathbb{R}^2$ is a row vector whit traspose $\mathbf{u}^t$, the function g_0 is such that g is locally integrable (Møller and Toftaker (2012)). Σ is a 2×2 symmetric positive definite matrix of the form $\Sigma = \omega^2 U_\theta \text{diag}(1, \zeta^2)U_\theta^t$ with ζ the anisotropy factor and

$$U_\theta = \begin{pmatrix} \cos(\theta) & -sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix},$$

see Gabriel (2013).

Spatio-temporal Inhibition process

Spatio-temporal Inhibition processes were presented by Gabriel et al. (2012), and either prevent (strict inhibition) or make unlikely the occurrence of

pairs of close events, resulting in patterns that are more regular in space and/or in time than a Poisson process of the same intensity. In a spatial simple sequential inhibition process (strict inhibition), let δ_s denote the minimum permissible distance between events and ρ_s the spatial intensity of the process. The proportion of the plane covered by non-overlapping discs of radius $\delta_\mathrm{s}/2$ is

$$\rho^* = \frac{\rho_\mathrm{s}\pi\delta_\mathrm{s}^2}{4},$$

which we call the packing density. The maximum achievable packing density is for a pattern of points in a regular triangular lattice at spacing δ_s, for which $\rho^* = \frac{1}{2}\sqrt{3}$. Depending on exactly how the points are generated, even this value of δs may not be feasible. Simple sequential inhibition processes in space and time are defined by the following algorithm. Consider a sequence of m events $(\mathbf{u}_i, s_i) \in W \times T$. Then,

i. $\mathbf{u}_1$ and s_1 are uniformly distributed in W and T respectively.

ii. At the k-th step of the algorithm, $k = 2, \ldots, m$, s_k is uniformly distributed on $W \cap \Delta_\mathrm{space}$ with

$$\Delta_\mathrm{space} = \{\mathbf{u} : \|\mathbf{u} - \mathbf{u}_j\| \geq \delta_\mathrm{s}, \quad j = 1, \ldots, k-1\}$$

and s_k is uniformly distributed on $T \cap \vartheta_\mathrm{time}$ with

$$\Delta_\mathrm{time} = \{s : |s - s_j| \geq \delta_t, \quad j = 1, \ldots, k-1\}.$$

To obtain a larger class of inhibition processes, we can extend condition ii. of the above algorithmic definition by introducing functions $p_\mathrm{s}(\mathbf{u})$ and $p_\mathrm{t}(s)$ that together determine the probability that a potential point at location $\mathbf{u}$ and time s will be accepted as a point of the process.

1.9 Non-parametric estimation

1.9.1 Estimation of intensity functions

Suppose we are given estimates of $\hat{\rho}_{\text{space}}$ and $\hat{\rho}_{\text{time}}$. These provide unbiased estimates of the expected number of observed points, i.e.

$$\int_W \hat{\rho}_{\text{space}}(\mathbf{u}) \, \mathrm{d}\mathbf{u} = \int_T \hat{\rho}_{\text{time}}(s) \, \mathrm{d}s = n.$$

Then the estimate of the spatio-temporal intensity function is given by

$$\hat{\rho}(\mathbf{u}, s) = \frac{\hat{\rho}_{\text{space}}(\mathbf{u}) \, \hat{\rho}_{\text{time}}(s)}{n}. \tag{1.46}$$

This provides an unbiased estimate of the expected number of observed points, i.e.

$$\int_{W \times T} \hat{\rho}(\mathbf{u}, s) \, \mathrm{d}(\mathbf{u}, s) = n,$$

For non-parametric estimation of the spatial intensity function, we follow Diggle (1985) and Berman and Diggle (1989) in using the kernel estimate as in the equation (1.28)

$$\hat{\rho}_{\text{space}}(\mathbf{u}) = \sum_{i=1}^{n} \frac{\kappa_\epsilon(\mathbf{u} - \mathbf{u}_i)}{c_{W,\epsilon}(\mathbf{u}_i)}, \qquad \mathbf{u} \in W, \tag{1.47}$$

where

$$\kappa_\epsilon(\mathbf{u}) = \frac{1}{\epsilon^2} \kappa\left(\frac{\mathbf{u}}{\epsilon}\right),$$

is a kernel with bandwidth $\epsilon > 0$, i.e, κ is a given density function.

Further,

$$c_{W,\epsilon}(\mathbf{u}_i) = \int_W \kappa_\epsilon(\mathbf{u} - \mathbf{u}_i) \, \mathrm{d}\mathbf{u},$$

is an edge-correction factor ensuring that $\int_W \hat{\rho}_{\text{space}}(\mathbf{u}) \, \mathrm{d}\mathbf{u} = n$.

A similar kernel estimate may be used for non-parametric estimation of $\rho_{\text{time}}(s)$ wihth $d = 1$, see Gabriel et al. (2012). Although these non-

parametric estimation procedures of the spatial and temporal intensity functions may only lead to approximately unbiased estimates, we will still use equation (1.46). In the literature of spatio-temporal point process it is common modelling the temporal intensity component using of parametric methods. For example, Gabriel and Diggle (2009) estimate the temporal intensity using a log-linear regression model. Exploratory spectral analysis of incidence times in their application suggested a marked annual and 4-monthly periodicity. Therefore, they fitted a harmonic regression model with 1-year period

$$\log \rho_{\text{time}}(t) = \delta_{d(t)} + \sum_{k=1}^{3} \{\alpha_k \cos(k\omega t) + \beta_k \sin(k\omega t)\} + \gamma t,$$

where $\omega = 2\pi/365$, γ denotes the trend and $d(t)$ identifies the day of the week for day $t = 1, \ldots, 1096$. See also Diggle (2013) and Tamayo-Uria et al. (2014).

1.9.2 Estimation of K-, g- and J-functions

Lemma 80 *Alternatively, an approximately unbiased non-parametric estimate of the K-function is given by*

$$\hat{K}(r,t) = \frac{1}{|W|\,|T|} \sum_{(\mathbf{u}_i,s_i),(\mathbf{u}_i,s_j)\in X}^{\neq} \frac{\mathbf{1}\left[\|\mathbf{u}_i - \mathbf{u}_j\| \leq r\right]\mathbf{1}\left[|s_i - s_j| \leq t\right]}{\hat{\rho}\left(\mathbf{u}_i, s_i\right)\hat{\rho}\left(\mathbf{u}_j, s_j\right)w^2\left(\mathbf{u}_i, \mathbf{u}_j\right)w^1\left(s_i, s_j\right)}$$

where $\sum^{\neq}$ means that the sum is over all pairs $(\mathbf{u}_i, s_i) \neq (\mathbf{u}_j, s_j)$ of the data points, w^2 and w^1 denote the Ripley's spatial edge-correction factors given by **Definition** *56.*

Note that in the temporal case $w^1(s_i, s_j) = 1$ if both ends of the interval of length $2|s_i - s_j|$ centred at s_i lie within T and $1/2$ otherwise (see Diggle et al. (1995)). For an unbiased estimator of the K-function see e.g., Gabriel (2013).

Non-parametric estimation of pair correlation functions are usually based on kernel methods (see Stoyan and Stoyan (1994) and Illian et al.

(2009)), where the specification of the bandwidth of the kernel is debatable, being this is the most problematic aspect of Baddeley et al. (2000).

Lemma 81 *The spatio-temporal pair correlation function defined in (1.45) can be estimated by*

$$\hat{g}(r,t) = \frac{1}{4\pi \, |W| \, |T| \, r} \sum_{(\mathbf{u}_i,s_i),(\mathbf{u}_i,s_j)\in X}^{\neq} \frac{\kappa_{1\epsilon}(\|\mathbf{u}_i - \mathbf{u}_j\| - r)\kappa_{2\delta}(|s_i - s_j| - t)}{\hat{\rho}\left(\mathbf{u}_i, s_i\right) \hat{\rho}\left(\mathbf{u}_j, s_j\right) w^2\left(\mathbf{u}_i, \mathbf{u}_j\right) w^1\left(s_i, s_j\right)},$$

(1.48)

where $\kappa_{1\epsilon}$ and $\kappa_{2\delta}$ are respectively one-dimensional kernel functions with bandwidths ϵ and δ.

For details see Gabriel et al. (2012) and Gabriel (2013).

Lemma 82 *For $r \geq 0$, $t \geq 0$, an estimator of $1 - G_{inhom}(r,t)$ is given by*

$$\sum_{(\mathbf{v},l)\in X\cap(W_{space}^{\ominus r} \times T_{time}^{\ominus t})} \left[\frac{\prod_{(\mathbf{u},s)\in(X\setminus\{(\mathbf{v},l)\})\cap((\mathbf{v},l)+S_{st})} \left(1 - \frac{\bar{\rho}}{\rho(\mathbf{u},s)}\right)}{|X \cap (W_{space}^{\ominus r} \times T_{time}^{\ominus t})|} \right]$$

where $W_{space}^{\ominus r} = \{\mathbf{u} \in W : d(\mathbf{u}, \partial W) \geq r\}$ and $T_{time}^{\ominus t} = \{s \in T : d(s, \partial T) \geq t\}$.

Lemma 83 *Given a finite point grid $L \subseteq W \times T$, the estimator of $1 - F_{inhom}(r,t)$ is given by*

$$\frac{1}{|L \cap (W_{space}^{\ominus r} \times T_{time}^{\ominus t})|} \sum_{(\mathbf{v},l)\in L\cap(W_{space}^{\ominus r} \times T_{time}^{\ominus t})} \left[\prod_{(\mathbf{u},s)\in X\cap((\mathbf{v},l)+S_{st})} \left(1 - \frac{\bar{\rho}}{\rho(\mathbf{u},s)}\right) \right]$$

The ratio of the estimates given in **Lemma 82** and **Lemma 83** gives an estimator of $J_{\text{inhom}}(r,t)$.

Second-order smoothing of spatial point patterns with small sample sizes: A family of kernel

Francisco J. Rodríguez-Cortés and **Jorge Mateu**,[1]

Department of Mathematics, Universitat Jaume I, Castellón, Spain

Abstract

We consider kernel-based non-parametric estimation of second-order product densities of spatial point patterns. We present a new family of optimal and positive kernels showing less variance than optimal kernels. This family generalises most of the classical and widely used kernel functions, such as Box or Epanechnikov kernels. We propose an alternative unbiased estimator for the product density function, and compare the performance of the estimator for several members of the family of optimal and positive kernels through MISE and relative efficiency. We present a simulation study to analyse the behaviour of such kernel functions, for three different spatial structures, for which we know the exact analytical form of the

---

[1]All authors contributed equally in this work.

product density, and under small sample sizes. Some known datasets are revisited.

Keywords: *Mean integrated square error, Kernel smoothing, Relative efficiency, Second-order product density function, Non-parametric estimation.*

2.1 Introduction

The first studies using smoothing techniques have been proposed in a wide range of applied sciences such as Actuary, Economy, Mathematical Finance, or Statistics, see e.g. Woolhouse (1870), De Forest (1873), Gram (1883), Macaulay (1931), and Hoem (1983). The idea of using the smoothing technique as a local regression method has its beginnings in the sixties, with kernel methods introduced in the estimation of the density by Akaike (1954), Parzen (1962), Rosenblatt (1971) and, in regression, by Nadaraya (1964) and Watson (1964). The problem of choosing a suitable kernel in the context of non-parametric estimation of curves has been deeply studied by Watson and Leadbetter (1963) through determining the kernel that minimises the mean integrated square error (MISE). Epanechnikov (1969) suggested a set of kernels which obtain the optimal convergence rate for the MISE within a class of densities. This work was later extended by Sacks and Ylvisaker (1981) to the multivariate case; these kernels are asymptotically optimal among non-negative kernels for two-differentiable densities. Müller and Gasser (1979), Gasser and Müller (1984) and Gasser et al. (1985) introduced two classes of optimal kernels that minimise the variance and the MISE, respectively. Falk (1983) and Mammitzsch (1984) studied in detail the problem of finding optimal kernels. Cline (1988) gave a new concept of admissible kernel density estimator when no other kernel estimator has uniformly smaller MISE using Fourier techniques. Messer and Goldstein (1993) introduced a new class of variable kernels depending on the smoothing parameter and having good MISE convergence properties, equivalent to smoothing splines. For more recent and detailed works see e.g. Fan and Marron (1994), Hart (1997), Efromovich (1999), and Berlinet and Thomas-Agnan (2004).

Point processes, which are random collections of points falling in some measurable space, have found use in describing an increasing number of naturally arising phenomena in a wide variety of applications, including epidemiology, ecology, forestry, mining, hydrology, astronomy, ecology, and meteorology (Daley and Vere-Jones (2003); Schoenberg et al. (2006)).

In the mid 20th century, interest spanned to spatial point processes, where each point represented the location of some object or event, such as a tree or sighting of a species (Cressie (1993); Illian et al. (2008)). The classical and reference model for spatial point processes is the Poisson process, where the number of points in any two disjoint sets are independent random variables; the name comes from the fact that for point processes with this independence property, the number of points in any measurable set follows a Poisson random variable (see Daley and Vere-Jones (2003)).

A spatial point pattern is a set of points $\{\mathbf{u}_i \in W : i = 1, \ldots, n\}$ for some bounded spatial region W. Very often, W is a sampling window within a much larger region and it is reasonable to regard the point pattern as a partial realisation of a stochastic planar point process, the events consisting of all points of the process which lie within W. The study of spatial point patterns has a long history in ecology and forestry (Goodall (1952); Pielou (1977); Ripley (1981); Comas et al. (2009); Comas and Mateu (2011)). Spatial point patterns have also found application in fields as diverse as cosmology Neyman and Scott (1958), archeology Hodder and Orton (1976), geography Cliff and Ord (1981), epidemiology Diggle and Richardson (1993), seismology Ogata (1998) or environmental sciences Juan et al. (2012). Recent textbooks related to the topic of analysis and modelling of point processes include Stoyan et al. (1987), Diggle (1983, 2003), Møller and Waagepetersen (2004), Baddeley et al. (2006), Illian et al. (2008) or Gelfand et al. (2010).

The connection between these two areas of research (kernel-based non-parametric estimation and spatial point processes) begins when Diggle (1983) analyses the use of a simple circular (spherical) kernel in the estimation of the intensity for one-dimensional point processes. Inference by non-parametric methods was introduced in the context of spatial point

processes by Diggle (1985), adapting the idea of Silverman (1978) to estimate the intensity function under the argument that $\widehat{\rho}(r)$ and $\widehat{f}(x)$ differ by a factor of n, because the intensity is a mean number of events per unit of area and, unlike the probability density function, does not integrate to one. Diggle (1985, 2003) discusses the choice of an optimal bandwidth for a specific model. In the same way that the intensity function has been analysed using non-parametric techniques, we can estimate second-order characteristics in spatial point processes such as the second-order product density function (see e.g. Akaike (1954); Brillinger (1975); Krickeberg (1982); Fiksel (1988b); Ohser and Mücklich (2000); Stoyan and Stoyan (2000)). In addition we can estimate the pair correlation function dividing the product density by the respective intensities (see e.g. Doguwa (1990); Ohser (1991); Stoyan and Stoyan (2000); Baddeley et al. (2000); Guan (2007); Comas et al. (2009); Juan et al. (2012)).

Cressie (1993), Stoyan and Stoyan (1994), Illian et al. (2008) claim that the choice of the kernel is not important and one should pay more attention to the choice of the bandwidth for the smoothing. But this fact should not be understood as if the choice of the kernel is irrelevant because the bandwidth depends on the kernel, and the accuracy of the estimate is not only a function of the bandwidth, but depends upon the pair kernel-bandwidth (κ, ϵ). In the literature of point processes it is commonly suggested the use of the Epanechnikov kernel arguing that it has shown certain optimality properties. However, previous publications do not give analytical arguments for this preference. Furthermore, the selection of the optimal bandwidth for this kernel is attributed to simulations and practical experience as $\epsilon = c/\sqrt{\rho}$, with $c \in [0.1, 0.2]$. The extended use of the Epanechnikov kernel in the spatial statistics literature is due to the fact that it is the optimum second-order kernel and asymptotically minimises the Mean Integrated Squared Error (MISE) and the Mean Square Error (MSE). The known theoretical arguments are all under asymptotic conditions. However it is quite often the case that in practice the number of points in the point pattern is not large enough (Diggle (1983, 2003); Comas et al. (2009); Comas and Mateu (2011)). And in this paper we focus

our interest over these cases. Our aim is to highlight the benefits of using alternative kernel functions showing in which cases they report a better performance than the widely used Epanechnikov kernel when estimating second-order product densities. In particular these new kernel functions will show a much better performance under cluster structures.

In Section 2.2 we present an alternative kernel family to the Epanechnikov kernel. In particular we use an optimal and positive kernel family with the aim to reduce the importance of the choice of the correct bandwidth. These kernels are composed by functions in $\mathbb{Q}[x]$, the space of rational functions. This family of positive kernels is also related to the optimal kernels. We consider a set of polynomials with a greater degree than those in Marron and Nolan (1988). In addition, this family of kernels has the property of including, as a particular case, the Epanechnikov kernel. In Section 2.3 we present a short review of the more relevant theoretical concepts on spatial point processes. We use an approximately unbiased estimator for the second-order product density function built as a combination between the Ripley's edge-corrector and the consideration in the denominator of pairwise distances between points of the process. In Section 2.4 we present a simulation study for a variety of scenarios usually encountered in the practice of point processes by comparing the behaviour of the new family with respect to the usual Epanechnikov kernel and other given optimum and positive kernels which also belong to the same family. We comment on choosing adequate bandwidth parameters. We compare the performance of several kernel functions through a measure of relative efficiency, and show that we obtain larger degrees of smoothing when using optimal and positive kernel functions of some high orders under cluster structures. An application to known datasets is also considered in Section 2.4. The paper ends with some conclusions.

2.2 An optimal and positive kernel family

2.2.1 Theoretical setup

In non-parametric statistics it is known that under regularity conditions for two optimality criteria (minimum variance, minimum MSE), including the ν-th order differentiability of the curve to be estimated, and adding the optimal smoothing parameter, the asymptotic optimal MISE of the kernel estimator for a density function is proportional to the functional $T(\kappa)^{2/(2k+1)}$, where

$$T(\kappa) = \left(\int \kappa^2(x) \mathrm{d}x \right)^k |M_k(\kappa)|, \qquad (2.1)$$

here $M_k(\kappa) = \int \kappa(x) x^k \mathrm{d}x$ and k is the order of the kernel κ, see Gasser and Rosenblatt (1979). This means that the estimator of the function is a particular case of the estimator of the derivative when the order is zero. The kernel function κ in (2.1) has order k, i.e. it satisfies the moment condition $\kappa \in \mathcal{K}_k$ with

$$\mathcal{K}_k = \left\{ \kappa \in L_2[-1,1] : M_j(\kappa) = 0, \text{ for } j = 1, \ldots, k-1 \text{ and } M_k(\kappa) \neq 0 \right\}, \qquad (2.2)$$

where L_2 is the space of square integrable functions. For the problem of minimising the functional $T(\kappa)$ over the set of functions $\mathcal{K}_k$, it is needed an additional condition. Indeed Gasser et al. (1985) imposed on κ the condition of minimum sign changes

$$\kappa \in \mathcal{N}_{k-2} = \left\{ \kappa \in L_2[-1,1] : \kappa \text{ has exactly } k-2 \text{ sign changes} \right\},$$

where $k-2$ is equal to the lower bound of the number of change signs required in order to satisfy (4.2), see Müller (1985). The variational problem to consider thus becomes

$$\min_{\kappa \in \mathcal{K}_k \cap \mathcal{N}_{k-2}} \left\{ \left(\int \kappa^2(x) \mathrm{d}x \right)^k |M_k(\kappa)| \right\}. \qquad (2.3)$$

The kernel functions that solve (2.3) are called optimal kernels. In non-parametric estimation within the point process literature, the so-called optimal kernels are commonly used by minimising the Asymptotic Mean Integrated Squared Error (AMISE). One of the most famous kernels that meets the property of being optimal is the Epanechnikov kernel. Throughout this paper we denote by $\kappa_O^{(k)}(x)$ the optimal kernel functions of order k. However, not all optimal kernels are useful for non-parametric estimation in point processes. If we consider a kernel function as a weighting function for point counts, it should assign positive values and then we are only interested in that subset of optimal kernels that are positive. The kernel with $k = 2$ is the kernel derived by Epanechnikov (1969).
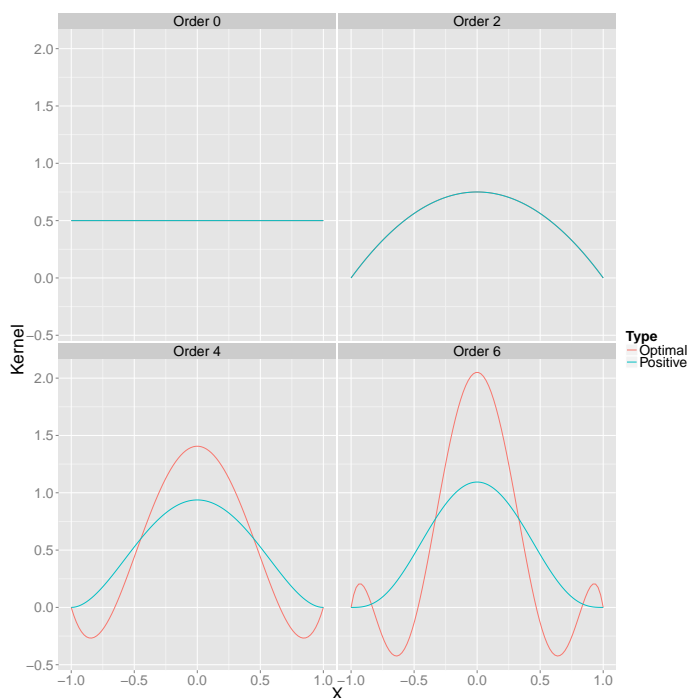
Figure 2.1: *Optimal and positive kernels under several orders.*

The canonical rescaling $\kappa_\epsilon(x) = \epsilon^{-1}\kappa(\epsilon^{-1}x)$ were the constant ϵ is called the bandwidth or smoothing parameter, and controls the amount of smoothing or local averaging, allows to separate the problems of kernel and bandwidth selection. It provides a fresh approach to the problem of optimal kernel selection, see Marron and Nolan (1988). They made it clear

that the best choice of the kernel does not depend on the knowledge of the bandwidth.

Our interest is on those kernels that are positive functions and that the corresponding weight diminishes as it moves away from the central point. We thus propose a new family of positive kernels and analyse their properties. Without losing generality, we add the condition that the kernel reaches its maximum at some $x = x_0$, and diminishes as it moves away from the center to the point where it has zero value at $x = -1$ and at $x = 1$, which implies positivity.

In order to carry this out, we use the family

$$\kappa_{\mathrm{P}}^{(\alpha,\beta)}(x) = C(1-x)^\alpha(1+x)^\beta, \tag{2.4}$$

in such a way that $\alpha + \beta = k$, which we refer to as the degree of the positive kernel, $\kappa_{\mathrm{P}}^{(\alpha,\beta)}(x) \approx O(x^{k+1})$, and C is such that

$$\int_{-1}^{1} \kappa_{\mathrm{P}}^{(\alpha,\beta)}(x)\mathrm{d}x = 1. \tag{2.5}$$

By making the change to the variable $y = (x+1)/2$ in (2.5), we obtain that

$$C = \frac{1}{2^{\alpha+\beta+1}\mathrm{Beta}(\alpha+1,\beta+1)}, \tag{2.6}$$

were $\mathrm{Beta}(\cdot,\cdot)$ is the beta function (see e.g. Luke (1969)). Using (**??**) and (2.6) it is possible to determine that the maximun of $\kappa_{\mathrm{P}}^{(\alpha,\beta)}(x)$ is found at $x = (\beta - \alpha)/(\alpha + \beta)$. In the case that $\alpha = \beta$, the family is symmetric and its maximun is found at $x = 0$. Therefore, for $\alpha = \beta = k/2$, the kernel family is given as

$$\kappa_{\mathrm{P}}^{(k)}(x) = \frac{\Gamma(k+2)}{2^{k+1}\Gamma\left(\frac{k}{2}+1\right)^2}(1-x^2)^{\frac{k}{2}}\mathbf{1}_{[-1,1]}(x) \tag{2.7}$$

were $\Gamma(\cdot)$ is the Gamma function (see e.g. Luke (1969)).

Naturally, the positive kernel family (2.7) satisfies the clasical require-

ments of non-parametric statistics, and has the properties

$$\int \kappa_{\mathrm{P}}^{(k)}(x)^2 \mathrm{d}x = \frac{\Gamma(k+2)^2 \Gamma(k+1)^2}{2\Gamma\left(\frac{k}{2}+1\right)^4 \Gamma(2k+2)}$$

and

$$\int \kappa_{\mathrm{P}}^{(k)}(x)x^k \mathrm{d}x = \frac{(1+(-1)^k)\Gamma(k+1)\Gamma(\frac{k}{2}+\frac{3}{2})}{2^{k+1}\Gamma(\frac{k}{2}+1)\Gamma(k+\frac{3}{2})},$$

which are useful in the calculation of the MISE for the estimated density function. Note that $\kappa_{\mathrm{P}}^{(k)}(x) \in \mathbb{Q}[x]$ the space of rational functions, furthermore for $k = 0$, the uniform kernel is obtained, $k = 2$ results in the Epachnenikov kernel, and in the case $k = 4$ the biweight kernel is obtained, as shown in Figure 2.1. These kernels are commonly used for their good performance in spatial smoothing, and they are both optimal and positive kernel functions.

2.2.2 Comparison of optimal and positive kernels

Asymptotic theory analyses kernel behaviour when the size of the sample $n \to \infty$, the bandwidth $\epsilon \to 0$, slowly enough for $n\epsilon \to \infty$ to be satisfied. However, in many applications the size of the sample is not large. The MSE and MISE depend on $\int \kappa(x)^2 \mathrm{d}x$ and $\int x^2 \kappa(x) \mathrm{d}x$. We now compare these quantities in both cases, for the optimal and positive kernels.

If k is even, we can write $k = 2n$ for $n \in \mathbb{N}$, and using the properties of the Gamma function we can rewrite the asymptotic bias in Gasser et al. (1985) as

$$\mathbb{B}(\kappa_{\mathrm{O}}^{(2n)}) = \frac{(-1)^{n+1}\Gamma(2n+2)\Gamma^2(2n+1)}{\Gamma^2(n+1)\Gamma(4n+2)}. \tag{2.8}$$

On the other hand, we have

$$\mathbb{B}(\kappa_{\mathrm{P}}^{(2n)}) = \frac{\Gamma(2n+1)\Gamma(n+\frac{3}{2})}{2^{2n}\Gamma(n+1)\Gamma(2n+\frac{3}{2})}. \tag{2.9}$$

Taking advantage of the properties of the Gamma function we have that

$$\Gamma(2n+2) = \frac{2^{2n+1}}{\sqrt{\pi}}\Gamma(n+1)\Gamma\left(n+\frac{3}{2}\right)$$

and

$$\Gamma(4n+2) = \frac{2^{4n+1}}{\sqrt{\pi}}\Gamma(2n+1)\Gamma\left(2n+\frac{3}{2}\right).$$

Replacing these into (2.8) and comparing with (2.9) we get that $\mathbb{B}(\kappa_{\mathrm{P}}^{(2n)}) = |\mathbb{B}(\kappa_{\mathrm{O}}^{(2n)})|$. Let us now analyse the kernel variances. The variance of the optimum kernel is given in Gasser et al. (1985), as

$$\mathbb{V}\mathrm{ar}_{\mathrm{O}} = \mathbb{V}\mathrm{ar}(\kappa_{\mathrm{O}}^{(k)}) = \frac{(k+1)k^2\Gamma^2(k+1)}{(2k+1)2^{2k}\Gamma^4(\frac{k}{2}+1)},$$

and the variance for the positive kernel is

$$\mathbb{V}\mathrm{ar}_{\mathrm{P}} = \mathbb{V}\mathrm{ar}(\kappa_{\mathrm{P}}^{(k)}) = \frac{\Gamma^2(k+2)\Gamma^2(k+1)}{2\Gamma^4(\frac{k}{2}+1)\Gamma(2k+2)}.$$

The ratio $\mathbb{V}\mathrm{ar}_{\mathrm{O}}/\mathbb{V}\mathrm{ar}_{\mathrm{P}}$ will be then

$$\frac{\mathbb{V}\mathrm{ar}_{\mathrm{O}}}{\mathbb{V}\mathrm{ar}_{\mathrm{P}}} = \frac{k^2\Gamma(2k+1)}{2^{2k-1}(k+1)\Gamma(k+1)^2}, \tag{2.10}$$

and using again the properties of the Gamma function in (2.10) we obtain

$$\frac{\mathbb{V}\mathrm{ar}_{\mathrm{O}}}{\mathbb{V}\mathrm{ar}_{\mathrm{P}}} = \frac{2k}{\sqrt{\pi}(k+1)}\frac{\Gamma\left(k+\frac{1}{2}\right)}{\Gamma(k)}. \tag{2.11}$$

The first factor is greater than 1 if $k \geq \frac{\sqrt{\pi}}{2-\sqrt{\pi}} \approx 7.8$, the second factor is always greater than 1, given that the $\Gamma(\cdot)$ function is an increasing monotone function. Then, from $k \geq 8$ even, we have that $\mathbb{V}\mathrm{ar}_{\mathrm{O}} > \mathbb{V}\mathrm{ar}_{\mathrm{P}}$, implying that for a kernel with an order larger than 8 can obtain a better smoothing (smoothing with a lower variance) for the same bandwidth if we rather use a positive kernel than an optimal one.

2.3   A second-order product density estimator for spatial point processes

A point process is a stochastic model governing the locations of events $\mathbf{u}$ in some set X. Because our interest is in spatial point processes, following Møller and Waagepetersen (2004), we formally consider a spatial point process with no multiple points as a random countable variable X. The events $\mathbf{u}$ of X are observed within a bounded spatial region $W \subset \mathbb{R}^2$, with area $|W| > 0$. $N(W)$ denotes the number of events falling inside of W. Assume that X has both intensity function $\rho(\cdot)$ and second-order product density function $\rho^{(2)}(\cdot)$, then for any non-negative Borel function f defined on $\mathbb{R}^2$, and any non-negative Borel function h defined on $\mathbb{R}^2 \times \mathbb{R}^2$,

$$\mathbb{E} \sum_{\mathbf{u} \in X} f(\mathbf{u}) = \int f(\mathbf{u}) \rho(\mathbf{u}) \mathrm{d}\mathbf{u}$$

and

$$\mathbb{E} \sum_{\mathbf{u}, \mathbf{v} \in X}^{\neq} h(\mathbf{u}, \mathbf{v}) = \int \int h(\mathbf{u}, \mathbf{v}) \rho^{(2)}(\mathbf{u}, \mathbf{v}) \, \mathrm{d}\mathbf{u} \, \mathrm{d}\mathbf{v},$$

where $\sum^{\neq}$ means that the summation goes over all pairs $(\mathbf{u}, \mathbf{v})$ with $\mathbf{u} \neq \mathbf{v}$.

The intensity function $\rho(\cdot)$ has an easy interpretation as the expected number of events per unit area. On the other hand, when describing variability and correlations of any pattern, we have to consider pairs of event, and the corresponding characteristics are called second-order measures such as the second-order product density function. A point process is said to be second-order weak stationary if $\rho(\mathbf{u}) = \rho$, a constant, and $\rho^{(2)}(\mathbf{u}, \mathbf{v}) = \rho^{(2)}(\|\mathbf{u} - \mathbf{v}\|)$, where $\| \cdot \|$ denotes the Euclidean norm. For infinitesimal $\mathrm{d}\mathbf{u}$ and $\mathrm{d}\mathbf{v}$, $\rho^{(2)}(\|\mathbf{u} - \mathbf{v}\|)\mathrm{d}\mathbf{u}\mathrm{d}\mathbf{v}$ can be interpreted as the probability that there is a point of the point process in each of two specified infinitesimal sets with areas $\mathrm{d}\mathbf{u}$ and $\mathrm{d}\mathbf{u}$ respectively. Under stationarity and isotropy, it is well known that $\rho^2 \mathrm{d}K(r) = 2\pi r \rho^{(2)}(r)\mathrm{d}r$, where $K(r)$ is the Ripley's K-function, see Ripley (1989). The second-order product density function is of interest as can be used to discriminate amongst sev-

eral spatial point structures. For example, for a spatial Poisson process $\rho^{(2)}(r) = \rho^2$. Values of the second-order product density function larger than ρ^2 indicate that the interpoint distances around r are relatively more frequent compared to those in a spatial Poisson process, which is typical of a spatial cluster process, and conversely, values of $\rho^{(2)}(r)$ smaller than ρ^2 indicate that the corresponding distances are rare and this is typical of a spatial inhibition process. The second-order product density function can take all values between zero and infinity, for large r it tends to ρ^2. The most typical spatial pattern models that have a closed analytical expression for the second-order product density function are the spatial modified Thomas process and the Matérn hard-core process, see Stoyan et al. (1987) and Illian et al. (2008). We focus on these processes to carry out the simulation study presented in Section 4.

In the context of the spatial cluster processes, the modified Thomas processes have the following closed form for the second-order product density function

$$\rho_{\mathrm{T}}^{(2)}(r; \rho_{\mathrm{p}}, \mu, \sigma) = \rho_{\mathrm{p}}^2 \mu^2 + \frac{\rho_{\mathrm{p}} \mu^2}{4\pi\sigma^2} \exp(-r^2/4\sigma^2), \qquad r \leq 0, \qquad (2.12)$$

where ρ_{p} is the intensity of the parents that follow a homogeneous Poisson process, the number of offspring per parent is Poisson with mean μ, and the distribution of the offspring around the parent is the symmetric normal with parameter σ. The intensity of a modified Thomas process is given by $\rho_{\mathrm{p}}\mu$.

In the inhibition process case, the Matérn hard-core process has the following expression for the second-order product density function

$$\rho_{\mathrm{M}}^{(2)}(r; \rho_{\mathrm{sp}}, r_0) = \begin{cases} 0, & r \leq r_0, \\[2mm] \frac{2\Gamma_{r_0}(r)(1-\exp(-\pi r_0^2 \rho_{\mathrm{sp}}))-2\pi r_0^2(1-\exp(-\rho_{\mathrm{sp}}\Gamma_{r_0}(r)))}{\pi r_0^2 \Gamma_{r_0}(r)(\Gamma_{r_0}(r)-\pi r_0^2)}, & r_0 < r \leq 2r_0, \\[2mm] \frac{1-\exp\left(-\pi r_0^2 \rho_{\mathrm{sp}}\right)}{\pi r_0^2}, & r > 2r_0, \end{cases}$$
$$(2.13)$$

where r_0 is the inhibition distance, ρ_{sp} is the intensity of a stationary Poisson process and

$$\Gamma_{r_0}(r) = 2\pi r_0^2 - 2r_0^2 \arccos\left(\frac{r}{2r_0}\right) + \frac{r}{2}\sqrt{4r_0^2 - r^2}, \qquad r_0 < r \leq 2r_0.$$

The intensity of the spatial Matérn hard-core process is $(1 - \exp\left(-\pi r_0^2 \rho_{\text{sp}}\right))$ $/\pi r_0^2$. These two point process models have been widely used in the literature and have been applied in a wide variety of practical scenarios and applications (Diggle (1983); Stoyan et al. (1987); Diggle (2003); Baddeley et al. (2006); Illian et al. (2008); Juan et al. (2012)). These processes are directly estimated and simulated from within the R library spatstat Baddeley and Turner (1995).

2.3.1 A non-parametric kernel estimator using the inter-event distance method

The empirical intensity and second-order product density functions are frequently estimated by non-parametric techniques such as kernel smoothing. Our attention is focused on the estimation of second-order characteristics, see Illian et al. (2008) for further details about aspects of estimating the intensity function. We use a similar estimator than Guan (2007), but we change the edge-correction to improve the estimation. Hence, for a stationary spatial point process, we have the following enhanced estimator for $\rho^{(2)}(\cdot)$

$$\widehat{\rho^{(2)}}_{\epsilon}(r) = \frac{1}{2\pi|W|} \sum_{\mathbf{u},\mathbf{v}\in X}^{\neq} \frac{\kappa_{\epsilon}(\|\mathbf{u}-\mathbf{v}\|-r)}{w(\mathbf{u},\mathbf{v})\|\mathbf{u}-\mathbf{v}\|}, \qquad r > \epsilon > 0, \qquad (2.14)$$

where κ_{ϵ} is the kernel function with bandwidth ϵ, and $w(\mathbf{u},\mathbf{v})$ is the isotropic edge-corrector proposed by Ripley (1989). One of the most important advantages of this estimator is given by the consideration in the denominator of the pairwise distances between points of the process $\|\mathbf{u}-\mathbf{v}\|$, and not the distances of a fine spatial grid r as has been suggested by many authors in the literature (Cressie (1993); Stoyan and Stoyan (1994); Diggle

(2003)). Consider for example a spatial cluster process with a high degree of aggregation. This implies that offsprings are very close with respect to the parents, therefore our estimator (2.14) prevents from an overestimation for small distances, as can be found when using the spatial grid. It is also noteworthy that the weights assigned by the edge-correction are based on the distance between pairs of the process and do not depend on the values of the spatial grid. An additional advantage of the estimator in (2.14) is that it better detects local spatial interaction present in the point pattern. Finally, this estimator is unbiased. Indeed, using Campbell's (Cressie (1993); Stoyan and Stoyan (1994)) and Fubini's theorem we have that

$$
\begin{aligned}
\mathbb{E}\left[\widehat{\rho^{(2)}}_\epsilon(r)\right] &= \frac{1}{2\pi|W|}\mathbb{E}\sum_{\mathbf{u},\mathbf{v}\in X}^{\neq}\frac{\kappa_\epsilon(\|\mathbf{u}-\mathbf{v}\|-r)}{w(\mathbf{u},\mathbf{v})\|\mathbf{u}-\mathbf{v}\|} \\
&= \frac{1}{2\pi|W|}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\frac{\kappa_\epsilon(\|\mathbf{x}-\mathbf{y}\|-r)}{w(\mathbf{x},\mathbf{y})\|\mathbf{x}-\mathbf{y}\|}\rho^{(2)}(\|\mathbf{x}-\mathbf{y}\|)\mathbf{1}_W(\mathbf{x})\mathbf{1}_W(\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y} \\
&= \frac{1}{2\pi|W|}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\frac{\kappa_\epsilon(\|\mathbf{z}\|-r)}{w(\mathbf{x},\mathbf{x}+\mathbf{z})\|\mathbf{z}\|}\rho^{(2)}(\|\mathbf{z}\|)\mathbf{1}_W(\mathbf{x})\mathbf{1}_W(\mathbf{x}+\mathbf{z})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{z} \\
&= \frac{1}{2\pi|W|}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\frac{\mathbf{1}_W(\mathbf{x})\mathbf{1}_W(\mathbf{x}+\mathbf{z})}{w(\mathbf{x},\mathbf{x}+\mathbf{z})}\mathrm{d}\mathbf{x}\frac{\kappa_\epsilon(\|\mathbf{z}\|-r)}{\|\mathbf{z}\|}\rho^{(2)}(\|\mathbf{z}\|)\mathrm{d}\mathbf{z} \\
&= \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty}\kappa_\epsilon(h-r)\rho^{(2)}(h)\mathrm{d}h\mathrm{d}\theta \\
&= \int_{-r/\epsilon}^{\infty}\kappa(s)\rho^{(2)}(r+\epsilon s)\mathrm{d}s.
\end{aligned}
$$

If r is a continuity point of $\rho^{(2)}(r)$, then

$$
\lim_{\epsilon\to 0}\mathbb{E}\left[\widehat{\rho^{(2)}}_\epsilon(r)\right] = \rho^{(2)}(r).
$$

Thus, when $\epsilon\to 0$, $\widehat{\rho^{(2)}}_\epsilon(r)$ is an approximately unbiased estimator for $\rho^{(2)}(r)$. Under a spatial Poisson process observed on any spatial window

W, it is easy to see that $\mathbb{E}\left[\widehat{\rho^{(2)}}_\epsilon(r)\right] = \rho^2$. We shall exploit this property through the MISE in the simulation study.

To obtain an estimator for the variance of the second-order product density function Stoyan et al. (1993a) and Stoyan and Stoyan (2000) proposed different methodologies using the set covariance function as edge-correction. However, our estimator (2.14) does not have the same structure as the quantity T in Ripley (1989), Stoyan et al. (1993a) and Daley and Vere-Jones (2003)), because $w(\mathbf{u}, \mathbf{v}) \neq w(\mathbf{v}, \mathbf{u})$, i.e. the edge-correction function is not symmetric implying that the term within the sum in (2.14) is not symmetric. This currently prevents from obtaining a close expression for the variance of (2.14).

2.4 Simulation study and data analysis

2.4.1 Performance of the estimator: a simulation study

We conduct a simulation study to show the behaviour of several kernel families in a variety of practical scenarios. In particular we want to highlight the better performance of some optimal and positive kernels in relation to the classical Epanechnikov kernel when the sample size of the point pattern is not large enough. The role of the kernel function comes implicit in the estimator (2.14) of the second-order product density. Considering that this estimator is unbiased, and knowing the closed analytical forms of the second-order product density function for the spatial modified Thomas process and the Matérn hard-core process (see equations (2.12) and (2.13)) we compare kernel performances through the mean integrated square error (MISE), which takes the form

$$\text{MISE}\left(\widehat{\rho^{(2)}}_\epsilon\right) = \mathbb{E}\left[\int_{r_{\inf} > \epsilon}^{r_{\sup}} \left(\widehat{\rho^{(2)}}_\epsilon(r) - \rho^{(2)}(r)\right)^2 \mathrm{d}r\right]. \qquad (2.15)$$

Note that for MISE computations it is highly important to rigorously define the lower limit of the integral as $r_{\inf} > \epsilon > 0$. The fact is that this function

counts in concentric rings with center $\mathbf{u}$ for each event $\mathbf{u} \in W$ under the condition $r - \epsilon \leq \|\mathbf{u} - \mathbf{v}\| \leq r + \epsilon$ for $\mathbf{u}, \mathbf{v} \in W$. If this lower limit condition is not respected, we can end up with overestimation and computational problems.

The main challenge with the MISE expression is that it depends on a complicated form depending on the bandwidth ϵ. This makes difficult to interpret the influence of the bandwidth on the performance of the second-order product density estimator via kernel. In order to show a more simplified interpretation and to purse a more fair comparison among the several simulation scenarios, we use instead

$$\text{RE}_k = \frac{1}{N_{\text{B}} N_{\text{sim}}} \sum_{\epsilon}^{N_{\text{B}}} \sum_{i=1}^{N_{\text{sim}}} \left(\text{MISE}_{ki\epsilon} - \overline{\overline{\text{MISE}}}_k \right)^2 \tag{2.16}$$

where N_{sim} and N_B are the number of pattern simulations and the number of selected bandwidths respectively, and

$$\overline{\overline{\text{MISE}}}_k = \frac{1}{N_{\text{B}} N_{\text{sim}}} \sum_{\epsilon}^{N_{\text{B}}} \sum_{i=1}^{N_{\text{sim}}} \text{MISE}_{ki\epsilon}.$$

The relative efficiency of the estimator is given by $(\min \{\text{RE}_k\} / \text{RE}_k) \times 100$.

The dependence of the estimator in (2.14) on the number of points, the spatial interaction structure and the combination kernel-bandwidth is clear. We thus examined the performance of the estimator for the second-order product density function via Monte Carlo simulation experiments over a range of scenarios. The second-order product density function was estimated using several kernel functions, and the corresponding MISE was calculated versus a sequence of bandwidths. We fixed the spatial window as $W = [0,1] \times [0,1]$, and simulated each of the three spatial point pattern structures (Poisson, modified Thomas cluster, and Matérn hard-core inhibition) with an expected number of points of $\mathbb{E}[N(W)] = n = 30, 50, 100$. We generated $N_{\text{sim}} = 1000$ realisations of the corresponding spatial point patterns for each particular scenario. We also considered a fine grid for the bandwidth ϵ spanning the sequence around the optimal bandwidth.

This optimal value was obtained using the dpik function in the kernsmooth package (Wand and Ripley (2013)) for the Poisson and modified Thomas cluster processes. Note that for the Matérn hard-core inhibition process it was necessary to impose the condition $\epsilon < r_{\text{inf}} < r_0$. The integral term in (2.15) was evaluated as a sum over a fine partition of the range of spatial distances, from $r_{\text{inf}} > \epsilon > 0$ to $r_{\text{sup}} = 0.25$ with small increments of spatial distances. We use of the library spatstat package (Baddeley and Turner (1995)) and connected Fortran subroutines to R.
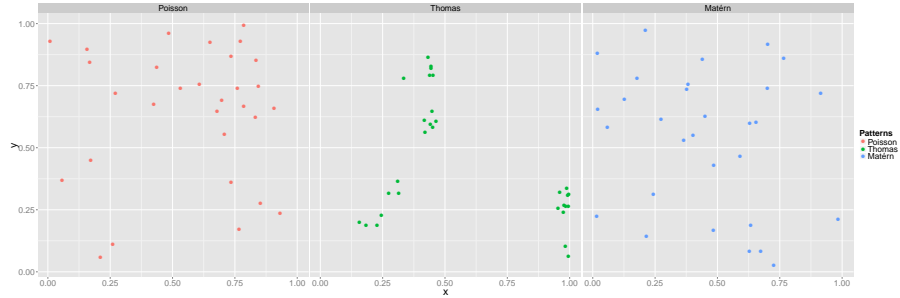
Figure 2.2: *Realisations of the spatial point patterns for $n = 30$.*
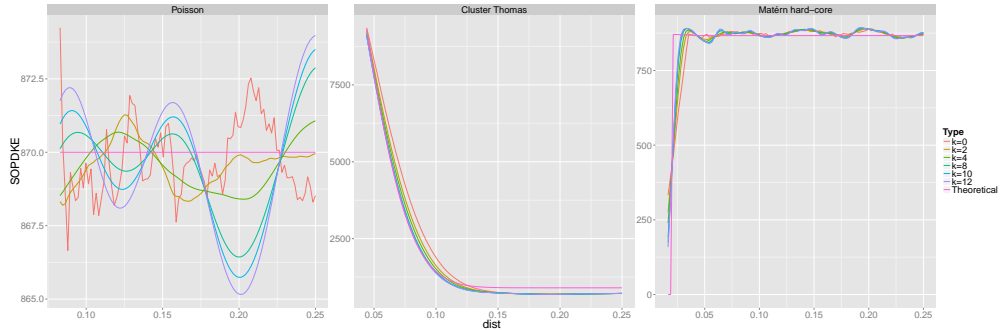
Figure 2.3: *Estimated second-order product density versus its theoretical form for $n = 30$.*

The point patterns were simulated as follows. For $n = 30$, the Poisson point patterns were generated with $\rho = 30$ and with an optimal bandwidth of $\epsilon = 0.0751912$. The modified Thomas cluster patterns were generated with $(\rho_{\text{p}}, \mu, \sigma) = (6, 5, 0.027)$ with an optimal bandwidth of $\epsilon = 0.04053382$. The Matérn hard-core inhibition patterns were generated

Figure 2.4: *MISE for the spatial point patterns with $n = 30$.*

with $(\rho_{\mathrm{sp}}, r_0) = (30, 0.02)$ and the optimal bandwidth was $\epsilon = 0.07513672$. For $n = 50$, the Poisson point patterns were generated with $\rho = 50$, with an optimal bandwidth of $\epsilon = 0.06053682$. The modified Thomas cluster patterns were generated with $(\rho_{\mathrm{p}}, \mu, \sigma) = (10, 5, 0.02)$ and an optimal bandwidth of $\epsilon = 0.03319133$. Finally, the Matérn hard-core inhibition patterns were generated with $(\rho_{\mathrm{sp}}, r_0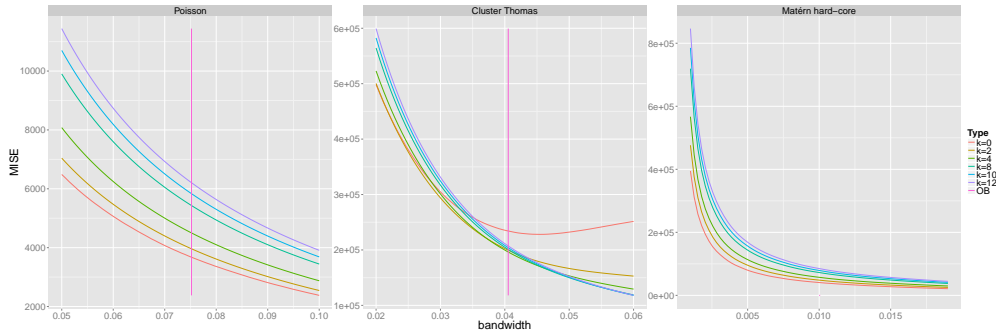) = (50, 0.014)$ and an optimal bandwidth of $\epsilon = 0.06116048$. For $n = 100$, the Poisson point patterns were generated with $\rho = 100$ and an optimal bandwidth of $\epsilon = 0.04524329$. The modified Thomas cluster patterns were generated with $(\rho_{\mathrm{p}}, \mu, \sigma) = (11, 90.04)$ and an optimal bandwidth of $\epsilon = 0.02652378$. The Matérn hard-core inhibition patterns were generated with $(\rho_{\mathrm{sp}}, r_0) = (100, 0.009)$ and an optimal bandwidth of $\epsilon = 0.04521666$.

Table 2.1 shows the relative efficiency of the overall MISE when estimating the second-order product density function under the three considered spatial structures, the three considered sample sizes, and using several orders (values of k) of the kernel functions. Note that the best performance is obtained for large values of the efficiency, which has a maximum value of 100. Note that our focus is on small sample sizes, thus we only use patterns with 100 points at most. Figures 2.2, 2.5 and 2.8 show particular realisations of these several types of point patterns, combining spatial structure and expected number of points.

Under the Poisson and Matérn hard-core inhibition cases (see Table 1) we note that the best performance is obtained with the Uniform or

Table 2.1: *Relative efficiency of MISE under spatial Poisson processes, modified Thomas cluster processes and Matérn hard-core inhibition processes.*

Pattern	Size	$k = 0$	$k = 2$	$k = 4$	$k = 8$	$k = 10$	$k = 12$
	$n = 30$	100	85.38382	74.06507	60.65028	56.18364	52.53986
Poisson	$n = 50$	100	87.06276	78.4698	68.03188	64.38115	61.30776
	$n = 100$	100	85.98112	74.6321	60.07972	55.05159	50.92702
	$n = 30$	84.59333	96.18569	99.13752	100	99.01781	99.01781
Cluster	$n = 50$	83.72175	98.07847	100	97.94181	96.25412	94.50055
	$n = 100$	97.78441	100	99.79232	98.35681	97.54602	96.73698
	$n = 30$	100	68.25449	47.84268	29.13356	24.18434	20.59619
Inhibition	$n = 50$	100	69.6394	48.63697	29.61517	24.66148	21.09242
	$n = 100$	100	70.08343	49.36589	30.36991	25.36599	21.74205

Box kernel, and efficiency decreases with the order of the kernel. However, when the spatial interaction in form of a cluster plays a role, things change. For the modified Thomas cluster processes shows that optimal and positive kernels with order larger or equal than $k = 8$ provide a better performance (larger efficiency) than the most simple kernel functions such as Box or Epanechnikov. This result is new in the literature and motivates the practical use of these new kernels.

Figures 2.3, 2.6 and 2.9 show the estimated second-order product density versus its theoretical form for the three sample sizes and spatial structures. It is clear that although they are not the most efficient, the optimal and positive kernels with the largest orders provide a larger degree of smoothing, in particular under Poisson structures.

Finally, Figures 2.4, 2.7 and 2.10 show the MISE values versus a range of bandwidths for the considered spatial point patterns. These Figures report similar results as the relative efficiencies in Table 1. Again, for cluster structures the larger order kernels provide lower MISE values. Note that in general MISE values decrease with the bandwidth. The vertical line in these Figures corresponds to the optimal bandwidth. In this case

Figure 2.5: *Realisations of the spatial point patterns for $n = 50$.*
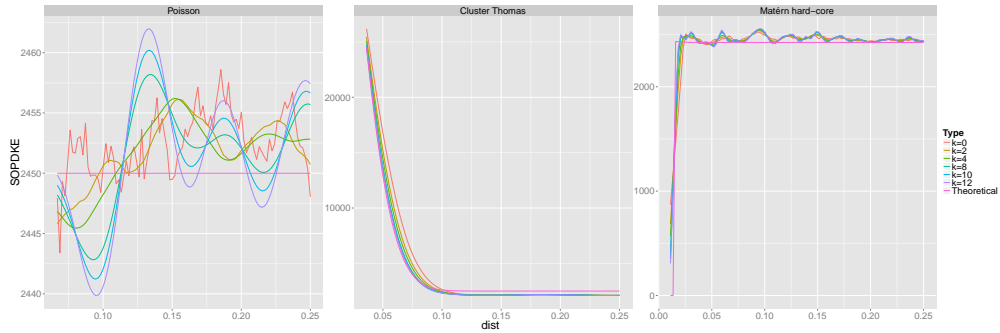
Figure 2.6: *Estimated second-order product density versus its theoretical form for $n = 50$.*
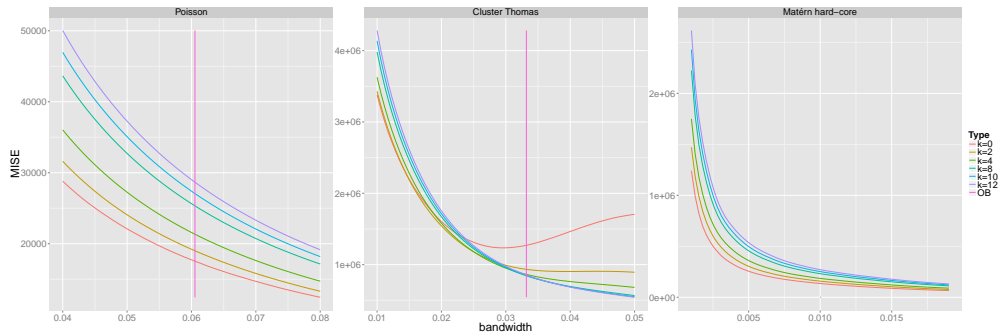
Figure 2.7: *MISE for the spatial point patterns with $n = 50$.*

it is also worth noting that there is not much difference, in terms of MISE and in terms of the several kernel functions, if we use a lower or larger bandwidth parameter than the optimal. In some sense we are safe against under- or over-estimation of the bandwidth.
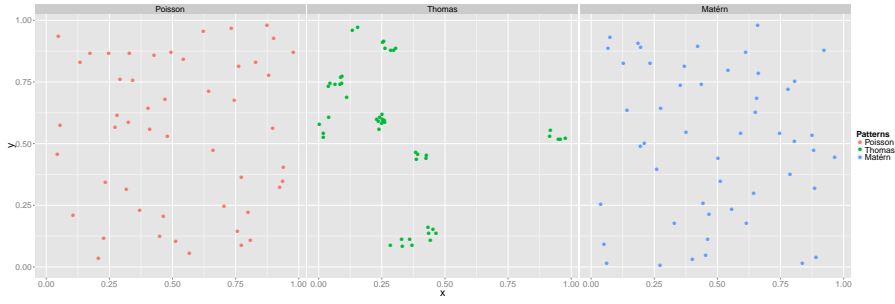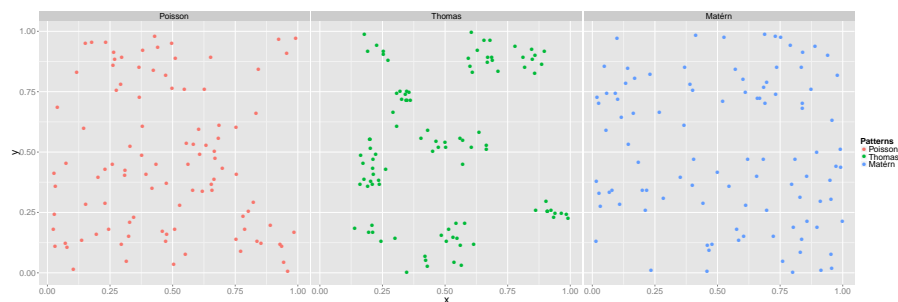
Figure 2.8: *Realisations of the spatial point patterns for $n = 100$.*
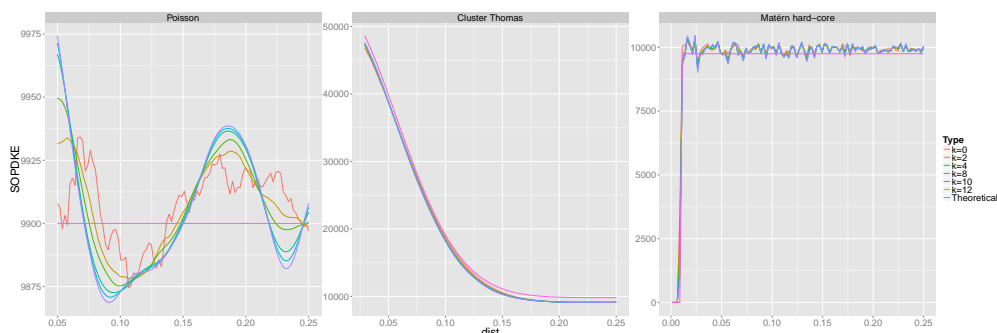
Figure 2.9: *Estimated second-order product density versus its theoretical form for $n = 100$.*
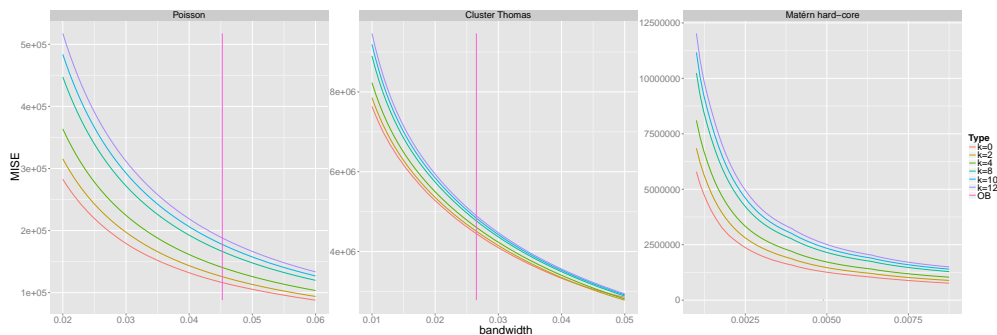
Figure 2.10: *MISE for the spatial point patterns with $n = 100$.*

2.4.2 Real-data analysis

We revisit the three examples of point patterns from Diggle (2003) as in Figure 2.11. No obvious second-order pattern appears for the Japanese pine saplings, but the redwood seedlings are clearly clustered, while the biological cell centers exhibit a regular or repulsion structure. A consistent

estimator of the second-order product density function can illuminate these features of the second-order structure of the point pattern X.
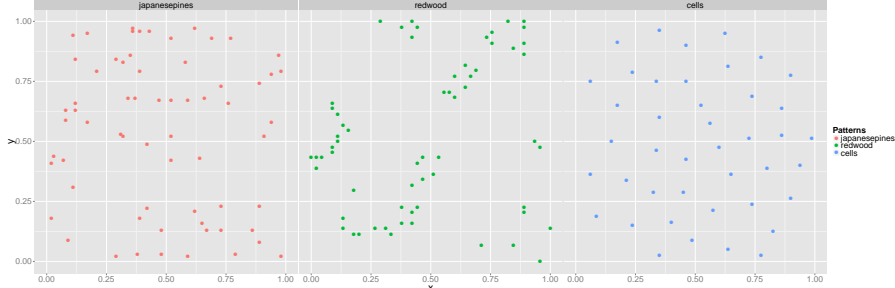
Figure 2.11: *Locations of 65 Japanese black pine saplings (left), 62 redwood seedlings (centre), and 42 biological cell centers (right) each observed on the unit square.*

Figure 2.12 presents estimates of the second-order product density function in order to illustrate the basic shape of the function. In all cases in Figure 2.12, ρ^2 stands for the estimated second-order product density function for a Poisson point process with the same number of points that the respective real dataset. The several second-order product density function estimates are consistent with the respective pattern structures. Japanese Pines can be considered a completely random point pattern since the estimate oscillates around $\hat{\rho}^2_{\text{japanese}} = 4160$ with $N(W) = 65$ and $\epsilon = 0.05328353$.
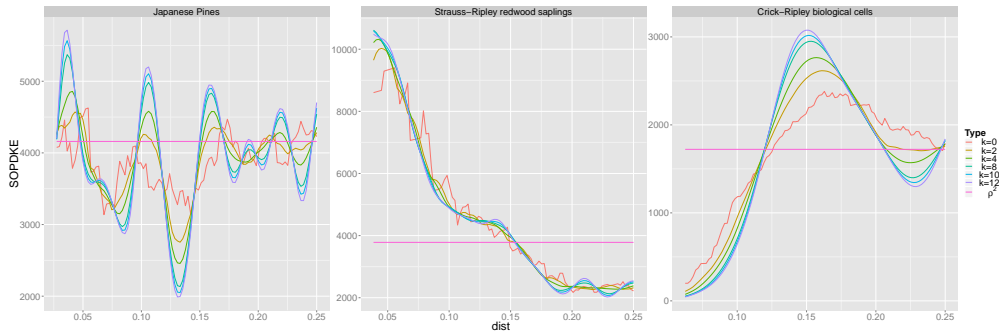
Figure 2.12: *Estimated second-order product density for the three selected point pattern datasets.*

This confirms the findings in the literature. The clustering behaviour in the redwdood seedling data is detected by a spike reporting values much

larger than $\hat{\rho}^2_{\text{redwood}} = 3782$ at small r with $N(W) = 62$ and $\epsilon = 0.04036764$. The regularity structure in the biological cells data is reflected in the low values of the product density compared with $\hat{\rho}^2_{\text{cells}} = 1722$ with $N(W) = 42$ and $\epsilon = 0.0568103$. In general we note again the greater degree of smoothing obtained with those kernels of higher order.

2.5 Conclusions

Estimation of characteristics of spatial point processes plays an important central role in the practice of point pattern analysis. This paper deals with kernel-based non-parametric estimation of second-order product densities. Estimation of product densities has not been paid much attention as estimation of first-order characteristics, despite they are crucial in detecting types of spatial interaction present in the point pattern. Literature has motivated the use of some simple kernels such as Box or Epanechnikov based on asymptotic arguments and extrapolating results from first-order characteristics to any order, and for any sample size. We have thus focused on those cases of small sample sizes, up to 100 events in the region of interest.

We have presented a new family of kernel functions based on optimal and positive kernel functions that generalises the widely used Box or Epanechnikov kernels. In addition we have built a new kernel-based estimator for the product density that is unbiased. We have shown that the performance of the new kernel functions depend upon the spatial structure present in the pattern data. In particular we suggest using optimal and positive kernel functions with order larger than eight when estimating second-order product densities under the presence of cluster structures. In these cases our proposed kernel functions provide both a better degree of smoothing and a closer behaviour to the theoretical product density. This is not a general property as under the presence of Poisson and regular structures the best efficiency is shown by the most simple kernel (the Box kernel). However, even in these cases, if we are more interested in smoothing, our new family of kernels with order larger than eight provide greater

degree of smoothing and thus the general pattern of the product density is better outlined. There are open lines of research coming out of our results. Under inhomogeneous point patterns, it would be of interest to analyse the behaviour of this kernel family and see if the presence of a trend changes the message found in this paper. An extension to the spatio-temporal case is also worth trying.

On the expected value and variance for an estimator of the spatio-temporal product density function

Francisco J. Rodríguez-Cortés[a], **Mohammad Ghorbani**[b],
Jorge Mateu[a], and **Dietrich Stoyan**[c],[1]

[a] Department of Mathematics, Universitat Jaume I, Castellón, Spain

[b] Department of Mathematics Sciences, Aalborg University, Aalborg, Denmark

[c] Institute of Statistics, TU Bergakademie Freiberg, Freiberg, Germany

Abstract

Second-order characteristics are used to analyse the spatio-temporal structure of the underlying point process, and thus these methods provide a natural starting point for the analysis of spatio-temporal point process data. We restrict our attention to the spatio-temporal product density

---

[1]All authors contributed equally in this work.

function, and develop a non-parametric edge-corrected kernel estimate of the product density under the second-order intensity-reweighted stationary hypothesis. First- and second-order spatio-temporal separability are introduced. The expectation and variance of the estimator are obtained, and close expressions derived under the Poisson case. A detailed simulation study is presented to compare our close expression for the variance with estimated ones for Poisson cases. The simulation experiments show that the theoretical form for the variance gives acceptable values, which can be used in practice. Finally, we apply the resulting estimator to data on the spatio-temporal distribution of invasive meningococcal disease in Germany.

Keywords and Phrases: *Point processes, Spatio-temporal separability, Second-order product density, Second-order intensity-reweighted stationarity, Variance*

3.1 Introduction

Spatial and spatio-temporal point patterns are increasingly available in a wide range of scientific settings, such as environmental sciences, climate prediction and meteorology, epidemiology, image analysis, agriculture and astronomy. Today, much attention is paid to spatio-temporal point processes, where each point represents the location and time of an event, and thus we have data of the form $(\mathbf{u}_i, s_i) \in W \times T \subset \mathbb{R}^2 \times \mathbb{R}$, $i = 1, ..., n$. There has been a lot of recent work on spatio-temporal models, and a variety of ad-hoc approaches have been suggested (Diggle (2006); Gabriel and Diggle (2009); Gelfand et al. (2010); Diggle (2013)). We consider here processes that are temporally continuous and either spatially continuous or spatially discrete on a sufficiently large support to justify formulating explicitly second-order spatio-temporal tools for the data.

For these processes second-order properties play an important role in the practical analysis of point patterns, in terms of exploratory and modelling strategies. Usually, the K-function and pair correlation function $(g(\cdot))$ are used for model checking (Møller and Ghorbani; 2013) and pa-

rameter estimation (Møller and Ghorbani; 2012), while the product density is used for explanatory analysis. The form of these functions helps to understand the type of interaction in the point pattern and to find suitable point process models.

In this context, separate analyses of the spatial and the temporal components are of limited value, because the scientific objectives of the analysis are to understand and to model the underlying spatio-temporally interacting stochastic mechanisms. There are basically two ways for modelling spatio-temporal point patterns (Diggle (2006); Daley and Vere-Jones (2008)). The first is descriptive and aims at providing an empirical description of the data, especially from second-order characteristics. The second is mechanistic and aims at constructing parametric point process models by specifying parametric models for the conditional intensity function. Here, we will consider the former and analyses will be based on extensions of the product density to summarize a spatio-temporal point pattern.

The inhomogeneous K-function has been extended to the spatio-temporal setting by Gabriel and Diggle (2009). Second-order characteristics are thus analysed from the spatio-temporal inhomogeneous K-function (STIK-function) or equivalently from the spatio-temporal pair correlation function under the assumption of second-order intensity re-weighted stationarity (Gabriel and Diggle (2009); Gabriel et al. (2010), Gabriel et al. (2012); Gabriel (2013)). Spatio-temporal separability of the STIK-function has been studied in Møller and Ghorbani (2012). These two functions rely very much upon first-order characteristics which are unknown in practice, and replacing the intensity by an estimate must be made carefully as it may imply bias (Baddeley et al. (2000); Gabriel (2013)). However, the product density does not show this problem, as will be shown in this paper.

Little attention has been paid so far to the first- and second-order moments (expected and variance values) of the second-order properties of spatio-temporal processes. And they are needed for performing statistical inference based on these characteristics. In the spatial context we can only refer to Ripley (1988) who developed variance expressions for a series of estimators of the spatial K-function for the Poisson process.

Then Stoyan et al. (1993b) approximated the variance of spatial product densities, and Cressie and Collins (2001a,b) obtained close expressions for the expected and variance values of the local spatial product densities. Nothing has been developed in the spatio-temporal context. In this paper we develop a non-parametric edge-corrected kernel estimate of the product density under the second-order intensity-reweighted stationary hypothesis. First- and second-order spatio-temporal separability are introduced. We extend the original ideas of Stoyan et al. (1993b) to the spatio-temporal case to develop exact and close expressions of the expectation and variance of the proposed estimator. Note that since estimated second-order characteristics deviate from their theoretical counterparts because of statistical fluctuations, error bounds for these functions are important. For example, they are needed to distinguish between statistical fluctuations in an estimated product density function and peaks which are due to real properties of the spatio-temporal point process under study.

Our estimator is accurate in estimating the spatio-temporal product density both under separable and non-separable cases. It is unbiased and we develop the close expression of its variance. The simulation experiments show that the formulae derived for this estimator give acceptable values, and thus can be used in practice.

The remainder of the paper is organised as follows. Section 3.2 provides a theoretical background on the first- and second-order properties of spatio-temporal point processes. In Section 3.3 we present the product density estimator and its expectation and variance for the general case, and under Poisson processes. Section 3.4 discusses the expectation and variance of the product density estimator under the hypothesis of spatio-temporal separability. We then present some simulation results in Section 3.5. Section 6 presents the analysis of the spatio-temporal distribution of invasive meningococcal disease in Germany. The paper ends with some final conclusions.

3.2 Definitions and backgrounds

Møller and Ghorbani (2012) discussed the second-order analysis of structured inhomogeneous spatio-temporal point processes. The definitions and notations introduced in that paper are used throughout the present paper. Following them, we consider a spatio-temporal point process with no multiple points as a random countable subset X of $\mathbb{R}^2 \times \mathbb{R}$, where a point $(\mathbf{u}, s) \in X$ corresponds to an event at $\mathbf{u} \in \mathbb{R}^2$ occurring at time $s \in \mathbb{R}$. In practice, we observe n events $\{(\mathbf{u}_i, s_i)\}$ of X within a bounded spatio-temporal region $W \times T \subset \mathbb{R}^2 \times \mathbb{R}$, with area $|W| > 0$, and with length $|T| > 0$. For formal definition of a point process based on measure theory see e.g. (Illian et al.; 2008; Chiu et al.; 2013).

For convenience, we introduce the following notations. Let $N(A)$ be the number of events falling in an arbitrary bounded region $A \subset W \times T$; $\Theta_n = \{(\mathbf{u}_1, s_1), \ldots, (\mathbf{u}_n, s_n) \in X\}$ be a set of n-tuples of events in X; $\int_{B^{\otimes k}} = \int_B \ldots \int_B$ for k times, where $B = W \times T$.

Assume that X has spatio-temporal nth-order product density function $\rho^{(n)}$, for $n \in \mathbb{N}$. For any non-negative Borel function f defined on $(\mathbb{R}^2 \times \mathbb{R})^{\otimes n}$,

$$
\mathbb{E} \sum_{\Theta_n}^{\neq} f((\mathbf{u}_1, s_1), \ldots, (\mathbf{u}_n, s_n)) = \int_{B^{\otimes n}} f((\mathbf{u}_1, s_1), \ldots, (\mathbf{u}_n, s_n))
$$
$$
\times \rho^{(n)}((\mathbf{u}_1, s_1), \ldots, (\mathbf{u}_n, s_n)) \, \mathrm{d}((\mathbf{u}_1, s_1), \ldots, (\mathbf{u}_n, s_n)), \qquad (3.1)
$$

where $\sum^{\neq}$ means that we sum over the n pairwise distinct points $(\mathbf{u}_1, s_1), \ldots, (\mathbf{u}_n, s_n)$ (see e.g. Illian et al. (2008); Chiu et al. (2013)).

3.2.1 First- and second-order properties

Considering (3.1), in particular for $n = 1$ and $n = 2$ the n-order product density function is respectively called the intensity function and the second-order product density function.

A process for which $\rho(\mathbf{u}, s) = \rho$ for all $(\mathbf{u}, s)$ is called homogeneous or

first-order stationary. Further, if $\rho^{(2)}((\mathbf{u}, s), (\mathbf{v}, l)) = \rho^{(2)}(\mathbf{u} - \mathbf{v}, s - l)$, the process is called second-order or weak stationary (Ghorbani; 2013).

3.2.2 Spatial and temporal components

It is assumed that the point process X is orderly, roughly meaning that coincident points cannot occur. That is, any pair of points $(\mathbf{u}, s)$ and $(\mathbf{v}, l)$ of X are distinct, so $\mathbf{u} \neq \mathbf{v}$ and $s \neq l$. We can therefore ignore the case where the spatial and temporal component processes X_{space} and X_{time} have multiple points, and following Møller and Ghorbani (2012) we define them by

$$X_{\text{space}} = \{\mathbf{u} : (\mathbf{u}, s) \in X, s \in T\}, \qquad X_{\text{time}} = \{s : (\mathbf{u}, s) \in X, \mathbf{u} \in W\}.$$

Note that, using this notation, it is clear that X_{space} depends on T, and X_{time} depends on W.

First-order properties

Assume that X has intensity function $\rho(\mathbf{u}, s)$, then

$$\rho_{\text{space}}(\mathbf{u}) = \int_T \rho(\mathbf{u}, s) \, \mathrm{d}s, \qquad \rho_{\text{time}}(s) = \int_W \rho(\mathbf{u}, s) \, \mathrm{d}\mathbf{u}.$$

Throughout the paper we assume first-order spatio-temporal separability, i.e.

$$\rho(\mathbf{u}, s) = \bar{\rho}_1(\mathbf{u})\bar{\rho}_2(s), \qquad (\mathbf{u}, s) \in \mathbb{R}^2 \times \mathbb{R}, \tag{3.2}$$

where $\bar{\rho}_1$ and $\bar{\rho}_2$ are non-negative functions. Considering the hypothesis of first-order spatio-temporal separability,

$$\rho(\mathbf{u}, s) = \frac{\rho_{\text{space}}(\mathbf{u})\rho_{\text{time}}(s)}{\int \rho(\mathbf{u}, s) \, \mathrm{d}(\mathbf{u}, s)}.$$

For stationary point process X, ρ, ρ_{space} and ρ_{time} are all constant. For non-parametric estimation of $\rho_{\text{space}}(\mathbf{u})$, $\rho_{\text{time}}(s)$ and $\rho(\mathbf{u}, s)$, (Ghorbani (2013)).

Second-order properties

Throughout the paper we assume that X is second-order intensity-reweighted stationary (SOIRS), i.e.

$$\rho^{(2)}((\mathbf{u}, s), (\mathbf{v}, l)) = \rho^{(2)}(\mathbf{u} - \mathbf{v}, s - l), \qquad (\mathbf{u}, s), (\mathbf{v}, l) \in \mathbb{R}^2 \times \mathbb{R} \quad (3.3)$$

(Baddeley et al.; 2000; Gabriel and Diggle; 2009; Gabriel; 2013). Further, if the process is isotropic, then $\rho^{(2)}(\mathbf{u} - \mathbf{v}, s - l) = \rho_0^{(2)}(\|\mathbf{u} - \mathbf{v}\|, |s - l|)$ for some non-negative function $\rho_0^{(2)}(\cdot)$, where $\|\cdot\|$ denotes the Euclidean distance in $\mathbb{R}^2$ and $|\cdot|$ denotes the usual distance in $\mathbb{R}$.

Using (3.1) (with $n = 2$) and (3.3) we obtain that X_{space} is SOIRS with second-order product density

$$\rho_{\text{space}}^{(2)}(\mathbf{u}, \mathbf{v}) = \rho_{\text{space}}^{(2)}(\mathbf{u} - \mathbf{v}) = \int_T \int_T \rho^{(2)}(\mathbf{u} - \mathbf{v}, s - l) \, \mathrm{d}s \, \mathrm{d}l. \qquad (3.4)$$

Analogously, X_{time} is SOIRS with

$$\rho_{\text{time}}^{(2)}(s, l) = \rho_{\text{time}}^{(2)}(s - l) = \int_W \int_W \rho^{(2)}(\mathbf{u} - \mathbf{v}, s - l) \, \mathrm{d}\mathbf{u} \, \mathrm{d}\mathbf{v}. \qquad (3.5)$$

It will always be clear from the context whether $\rho_{\text{space}}^{(2)}$ is considered to be a function defined on $\mathbb{R}^2 \times \mathbb{R}^2$ or $\mathbb{R}^2$, and whether $\rho_{\text{time}}^{(2)}$ is considered to be a function defined on $\mathbb{R} \times \mathbb{R}$ or $\mathbb{R}$.

Spatio-temporal separability of the product density function

The spatio-temporal product density function is separable if

$$\rho^{(2)}((\mathbf{u}, s), (\mathbf{v}, l)) = \bar{\rho}_1^{(2)}(\mathbf{u}, \mathbf{v}) \bar{\rho}_2^{(2)}(s, l)$$

for non-negative functions $\bar{\rho}_1^{(2)}$ and $\bar{\rho}_2^{(2)}$. Under the assumption (3.3) of SOIRS, this hypothesis can be rewritten as

$$\rho^{(2)}(\mathbf{u} - \mathbf{v}, s - l) = \bar{\rho}_1^{(2)}(\mathbf{u} - \mathbf{v}) \bar{\rho}_2^{(2)}(s - l), \qquad (\mathbf{u}, s), (\mathbf{v}, l) \in \mathbb{R}^2 \times \mathbb{R}. \quad (3.6)$$

Considering (3.4), (3.5), and (3.6),

$$\rho_{\text{space}}^{(2)}(\mathbf{u} - \mathbf{v}) = \bar{\rho}_1^{(2)}(\mathbf{u} - \mathbf{v}) \int_T \int_T \bar{\rho}_2^{(2)}(s - l) \, \mathrm{d}s \, \mathrm{d}l, \qquad (3.7)$$

and

$$\rho_{\text{time}}^{(2)}(s - l) = \bar{\rho}_2^{(2)}(s - l) \int_W \int_W \bar{\rho}_1^{(2)}(\mathbf{u} - \mathbf{v}) \, \mathrm{d}\mathbf{u} \, \mathrm{d}\mathbf{v}. \qquad (3.8)$$

By substituting (3.7) and (3.8) in (3.6),

$$\rho^{(2)}(\mathbf{u} - \mathbf{v}, s - l) = \frac{\rho_{\text{space}}^{(2)}(\mathbf{u} - \mathbf{v})\rho_{\text{time}}^{(2)}(s - l)}{\int \int \rho^{(2)}(\mathbf{u} - \mathbf{v}, s - l) \, \mathrm{d}(\mathbf{u}, s) \, \mathrm{d}(\mathbf{v}, l)}. \qquad (3.9)$$

As in the spatio-temporal first-order case, equation (3.9) suggests that

$$\rho^{(2)}((\mathbf{u}, s), (\mathbf{v}, l)) \propto \rho_{\text{space}}^{(2)}(\mathbf{u}, \mathbf{v})\rho_{\text{time}}^{(2)}(s, l).$$

Suppose we are given estimates $\widehat{\rho^{(2)}}_{\text{space}}(\mathbf{u} - \mathbf{v})$ and $\widehat{\rho^{(2)}}_{\text{time}}(s - l)$. If these are unbiased estimates of the expected number of distinct pairs of events, i.e. $\int_W \widehat{\rho^{(2)}}_{\text{space}}(\mathbf{u} - \mathbf{v}) \, \mathrm{d}\mathbf{u} \, \mathrm{d}\mathbf{v} = \int_T \widehat{\rho^{(2)}}_{\text{time}}(s - l) \, \mathrm{d}s \, \mathrm{d}l = n(n - 1)$, then the estimate of the spatio-temporal product density function given by

$$\widehat{\rho^{(2)}}(\mathbf{u} - \mathbf{v}, s - l) = \frac{\widehat{\rho^{(2)}}_{\text{space}}(\mathbf{u} - \mathbf{v})\widehat{\rho^{(2)}}_{\text{time}}(s - l)}{n(n - 1)},$$

is also a ratio unbiased estimate of the expected number of observed points. See more details in Section 3.3.

3.2.3 Relationship between the product density and the K-function

For a SOIRS, isotropic, spatio-temporal point process X, Gabriel and Diggle (2009) extended the inhomogeneous K-function from the spatial to the spatio-temporal case. They defined the spatio-temporal inhomogeneous

K-function as

$$K(r,t) = \int \mathbf{1}\left[\|\mathbf{u}\| \le r, |s| \le t\right] g_0(\mathbf{u}, s) \mathrm{d}(\mathbf{u}, s), \qquad r > 0, \quad t > 0, \quad (3.10)$$

where $\mathbf{1}[\cdot]$ denotes the indicator function, and $g_0(\mathbf{u}, s)$ (with the abuse of the notations $\mathbf{u}$ and s for $\mathbf{u} = \|\mathbf{u} - \mathbf{v}\|$ and $s = |s - l|$) is the spatio-temporal pair correlation function. For a Poisson process, $g_0 = 1$ and $K(r,t) = 2\pi r^2 t$. For an unbiased estimator of the K-function, see Gabriel (2013).

Considering the hypothesis of the first- and second-order spatio-temporal separabilities, for isotropic point process X and for non-negative Borel functions $\bar{K}_1$ and $\bar{K}_2$,

$$K(r,t) = \bar{K}_1(r)\bar{K}_2(t), \qquad r > 0, \quad t > 0. \qquad (3.11)$$

Assume that X is isotropic, and X_{space} and X_{time} have pair correlation functions g_{space} and g_{time} respectively. The corresponding spatial and temporal K-functions are

$$K_{\text{space}}(r) = \int_{\|\mathbf{u}\| \le r} g_{\text{space}}(\mathbf{u}) \, \mathrm{d}\mathbf{u}, \qquad r > 0,$$

and

$$K_{\text{time}}(t) = \int_{-t}^{t} g_{\text{time}}(s) \, \mathrm{d}s, \qquad t > 0.$$

Both in the stationary and isotropic case, and in the SOIRS and isotropic case, the spatio-temporal pair correlation function is proportional to the derivative of $K(r,t)$ with respect to r and t. So, in the planar case using (3.10),

$$g_0(r,t) = \frac{1}{4\pi r} \frac{\partial^2 K(r,t)}{\partial r \partial t}.$$

Thus, for the SOIRS and isotropic point process X,

$$\rho^{(2)}((\mathbf{u}, s), (\mathbf{v}, l)) = \frac{\rho(\mathbf{u}, s)\rho(\mathbf{v}, l)}{4\pi r} \frac{\partial^2 K(r,t)}{\partial r \partial t}, \qquad r > 0, \quad t > 0.$$

Further, under spatio-temporal separability (3.2) and (3.11), we have that

$$\rho^{(2)}(r,t) = \frac{1}{c_{\mathrm{SO}}} \left(\frac{\rho_{\mathrm{space}}(\mathbf{u})\rho_{\mathrm{space}}(\mathbf{v})}{2\pi r} \frac{\partial K_{\mathrm{space}}(r)}{\partial r} \right) \left(\frac{\rho_{\mathrm{time}}(s)\rho_{\mathrm{time}}(l)}{2} \frac{\partial K_{\mathrm{time}}(t)}{\partial t} \right).$$

(3.12)

Here

$$c_{\mathrm{SO}} = c_{\mathrm{SO}}^s \times c_{\mathrm{SO}}^t,$$

with

$$c_{\mathrm{SO}}^s = \left(\int_W \bar{\rho}^{(2)}(\mathbf{u})\,\mathrm{d}\mathbf{u} \right) \qquad \text{and} \qquad c_{\mathrm{SO}}^t = \left(\int_T \bar{\rho}^{(2)}(s)\,\mathrm{d}s \right),$$

and then

$$c_{\mathrm{SO}} = \left(\iint \rho^{(2)}(\mathbf{u},s)\,\mathrm{d}(\mathbf{u},s) \right),$$

which can be approximated by $n(n-1)$. Hence

$$\rho_{\mathrm{space}}^{(2)}(r) \propto \frac{\rho_{\mathrm{space}}(\mathbf{u})\rho_{\mathrm{space}}(\mathbf{v})}{2\pi r} \frac{\partial K_{\mathrm{space}}(r)}{\partial r}, \ \rho_{\mathrm{time}}^{(2)}(t) \propto \frac{\rho_{\mathrm{time}}(s)\rho_{\mathrm{time}}(l)}{2} \frac{\partial K_{\mathrm{time}}(t)}{\partial t}.$$

For a stationary and isotropic point process X,

$$\rho^{(2)}(r,t) = \frac{\rho^2}{4\pi r} \frac{\partial^2 K(r,t)}{\partial r \partial t}.$$

(3.13)

Moreover,

$$\rho_{\mathrm{space}}^{(2)}(r) \propto \frac{\rho_{\mathrm{space}}^2}{2\pi r} \frac{\partial K_{\mathrm{space}}(r)}{\partial r}, \qquad \rho_{\mathrm{time}}^{(2)}(t) \propto \frac{\rho_{\mathrm{time}}^2}{2} \frac{\partial K_{\mathrm{time}}(t)}{\partial t}.$$

3.3 Estimation of the product density

We avoid estimating the product density by applying numerical differentiation to an estimate of $\rho^2 K(r,t)$. Alternatively, considering that $\rho^2 K(r,t)$ stands for the expected number of ordered pairs of distinct points per unit area of the observation window with pairwise distance and time lag less than r and t, by extending the idea in Stoyan (1987) and Stoyan et al.

(1995), we directly estimate the product density using a non-parametric edge-corrected kernel estimate.

A spatio-temporal kernel density estimate of $\rho^2 \partial K(r,t)/\partial r \partial t$ takes the basic form of a smoothed three-dimensional histogram,

$$(|W||T|)^{-1} \sum_{(\mathbf{u},s),(\mathbf{v},l)\in X}^{\neq} \kappa_{\epsilon\delta}(\|\mathbf{u}-\mathbf{v}\| - r, |s-l| - t).$$

We assume that the kernel function $\kappa_{\epsilon\delta}(\cdot,\cdot)$ has the multiplicative form

$$\kappa_{\epsilon\delta}(\|\mathbf{u}-\mathbf{v}\| - r, |s-l| - t) = \kappa_{1\epsilon}\left(\|\mathbf{u}-\mathbf{v}\| - r\right) \kappa_{2\delta}\left(|s-l| - t\right),$$

where $\kappa_{2\delta}$ and $\kappa_{1\epsilon}$ are respectively one-dimensional kernel functions with bandwidths ϵ and δ.

(a) Counting form in the spatial case.   (b) Counting form in the spatio-temporal case.
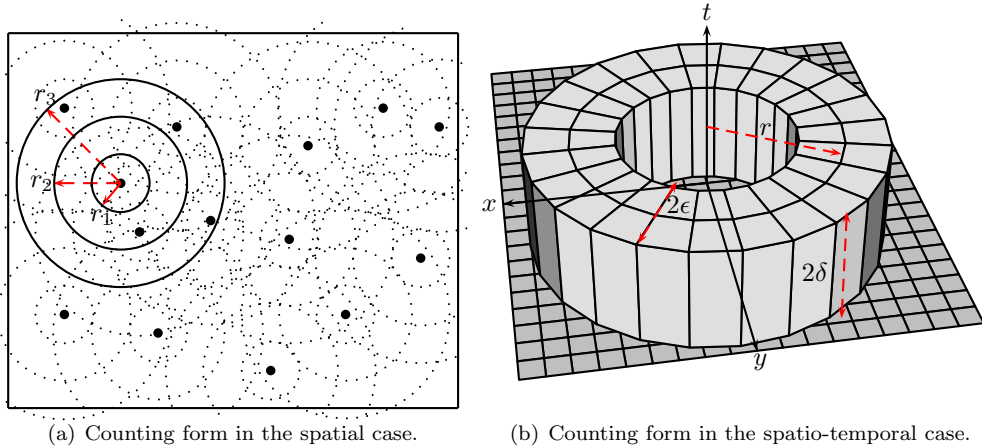
Figure 3.1: *Counting method for assigning weights in the estimation.*

Figure (3.1) illustrates how to count through cylindrical shells whose shape and volume change depending on the values of r and t; setting the spatial and temporal bandwidths, we define the weights for the global estimate. Furthermore, r and t are defined by empirical step functions in order to build a mesh that fits the nature of the data to obtain a enough good response surface to the interpretation of the phenomenon.

By extending the idea in Ohser (1983), an edge-corrected kernel estimate of the product density function (3.13) is given by

$$\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t) = \sum_{(\mathbf{u},s),(\mathbf{v},l)\in X}^{\neq} \frac{\kappa_{1\epsilon}\left(\|\mathbf{u}-\mathbf{v}\|-r\right)\kappa_{2\delta}\left(|s-l|-t\right)}{4\pi r\gamma_W(r)\gamma_T(t)}, \qquad (3.14)$$

with $r > \epsilon > 0$, $t > \delta > 0$. Here $\gamma_W(r)$ and $\gamma_T(t)$ are the spatial and temporal set covariance functions, respectively. For a convex region W, a general approximation formula for $\gamma_W(r)$ for small r is given by

$$\gamma_W(r) \approx |W| - \frac{U(W)}{\pi}r,$$

where $U(W)$ is the perimeter of W, and for a small t, $\gamma_T(t) = |T| - t$.

Under the hypothesis of spatio-temporal separability, and considering (3.9),

$$\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t) \simeq \frac{\widehat{\rho^{(2)}}_{\text{space},\epsilon}(r)\widehat{\rho^{(2)}}_{\text{time},\delta}(t)}{n(n-1)}, \qquad (3.15)$$

with

$$\widehat{\rho^{(2)}}_{\text{space},\epsilon}(r) = \sum_{\mathbf{u},\mathbf{v}\in X_{\text{space}}}^{\neq} \frac{\kappa_{1\epsilon}(\|\mathbf{u}-\mathbf{v}\|-r)}{2\pi r\gamma_W(r)}, \qquad r > \epsilon > 0,$$

and

$$\widehat{\rho^{(2)}}_{\text{time},\delta}(t) = \sum_{s,l\in X_{\text{time}}}^{\neq} \frac{\kappa_{2\delta}(|s-l|-t)}{2\gamma_T(t)}, \qquad t > \delta > 0.$$

3.3.1 Expectation and variance of the product density

In this section the expectation and variance of the product density estimator (3.14) is obtained by considering the general case. The corresponding moments of the product density estimator under the hypothesis of separability are developed in Section 3.4.

Expectation

Using (3.1) with $n = 2$, the estimator (3.14) satisfies

$$\mathbb{E}\left[\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t)\right] = \int\limits_{-r/\epsilon}^{\infty}\int\limits_{-t/\delta}^{\infty}\frac{\kappa_1(u)\kappa_2(v)\gamma_W(r+\epsilon u)\gamma_T(t+\delta v)}{r\gamma_W(r)\gamma_T(t)} \tag{3.16}$$

$$\times \rho^{(2)}(r+\epsilon u,t+\delta v)(r+\epsilon u)\mathrm{d}u\,\mathrm{d}v$$

The detailed proof is as follows. Applying formula (3.1), the Fubini's theorem and some elementary changes of variables in (3.14), we have that

$$\mathbb{E}\sum_{(\mathbf{u},s),(\mathbf{v},l)\in X}^{\neq}\frac{\mathbf{1}_W(\mathbf{u})\mathbf{1}_W(\mathbf{v})\mathbf{1}_T(s)\mathbf{1}_T(l)\kappa_{1\epsilon}\left(\|\mathbf{u}-\mathbf{v}\|-r\right)\kappa_{2\delta}\left(|s-l|-t\right)}{4\pi r\gamma_W(r)\gamma_T(t)}$$

$$=\int\limits_{W\times T}\int\limits_{W\times T}\frac{\kappa_{1\epsilon}(\|\mathbf{x}-\mathbf{y}\|-r)\kappa_{2\delta}(|\xi-\eta|-t)}{4\pi r\gamma_W(r)\gamma_T(t)}$$

$$\times \rho^{(2)}(\|\mathbf{x}-\mathbf{y}\|,|\xi-\eta|)\,\mathrm{d}(\mathbf{x},\xi)\,\mathrm{d}(\mathbf{y},\eta)$$

$$=\int\limits_{(W-h_1)\times(T-h_2)}\int\limits_{W\times T}\frac{\kappa_{1\epsilon}(\|\mathbf{h_1}\|-r)\kappa_{2\delta}(|h_2|-t)}{4\pi r\gamma_W(r)\gamma_T(t)}$$

$$\times \rho^{(2)}(\|\mathbf{h_1}\|,|h_2|)\,\mathrm{d}(\mathbf{h_1},h_2)\,\mathrm{d}(\mathbf{y},\eta)$$

$$=\int\limits_{\mathbb{R}^2\times\mathbb{R}}\frac{\kappa_{1\epsilon}(\|\mathbf{h_1}\|-r)\kappa_{2\delta}(|h_2|-t)\gamma_W(h_1)\gamma_T(h_2)}{4\pi r\gamma_W(r)\gamma_T(t)}\rho^{(2)}(\|\mathbf{h_1}\|,|h_2|)\,\mathrm{d}(\mathbf{h_1},h_2)$$

$$=\int\limits_0^{\infty}\int\limits_{\mathbb{R}}\frac{\kappa_{1\epsilon}(R-r)\kappa_{2\delta}(|h_2|-t)\gamma_W(R)\gamma_T(h_2)}{2r\gamma_W(r)\gamma_T(t)}\rho^{(2)}(R,|h_2|)\,R\,\mathrm{d}R\,\mathrm{d}h_2$$

$$=\int\limits_{-r/\epsilon}^{\infty}\int\limits_0^{\infty}\frac{\kappa_1(u)\kappa_2((h_2-t)/\delta)\gamma_W(r+u\epsilon)\gamma_T(h_2)}{2\delta r\gamma_W(r)\gamma_T(t)}$$

$$\times \rho^{(2)}(r+u\epsilon,h_2)(r+u\epsilon)\,\mathrm{d}u\,\mathrm{d}h_2$$

$$+\int\limits_{-r/\epsilon}^{\infty}\int\limits_{-\infty}^{0}\frac{\kappa_1(u)\kappa_2((-h_2-t)/\delta)\gamma_W(r+u\epsilon)\gamma_T(h_2)}{2\delta r\gamma_W(r)\gamma_T(t)}$$

$$\times \rho^{(2)}(r+u\epsilon,-h_2)(r+u\epsilon)\mathrm{d}u\,\mathrm{d}h_2$$

$$= \int\limits_{-r/\epsilon}^{\infty} \int\limits_{-t/\delta}^{\infty} \frac{\kappa_1(u)\kappa_2(v_1)\gamma_W(r+u\epsilon)\gamma_T(\delta v_1 + t)}{2r\gamma_W(r)\gamma_T(t)}$$

$$\times \rho^{(2)}(r+u\epsilon, \delta v_1 + t)(r+u\epsilon)\mathrm{d}u\,\mathrm{d}v_1$$

$$+ \int\limits_{-r/\epsilon}^{\infty} \int\limits_{-t/\delta}^{\infty} \frac{\kappa_1(u)\kappa_2(v_2)\gamma_W(r+u\epsilon)\gamma_T(\delta v_2 + t)}{2r\gamma_W(r)\gamma_T(t)}$$

$$\times \rho^{(2)}(r+u\epsilon, \delta v_2 + t)(r+u\epsilon)\mathrm{d}u\,\mathrm{d}v_2$$

$$= \int\limits_{-r/\epsilon}^{\infty} \int\limits_{-t/\delta}^{\infty} \frac{\kappa_1(u)\kappa_2(v)\gamma_W(r+u\epsilon)\gamma_T(\delta v + t)}{r\gamma_W(r)\gamma_T(t)}$$

$$\times \rho^{(2)}(r+u\epsilon, \delta v + t)(r+u\epsilon)\mathrm{d}u\,\mathrm{d}v,$$

which shows the result.

If (r,t) is a continuity point of $\rho^{(2)}(r,t)$, then

$$\lim_{(\epsilon,\delta)\to(0,0)} \mathbb{E}\left[\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t)\right] = \rho^{(2)}(r,t).$$

Hence, $\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t)$ is an approximately unbiased estimator for the spatio-temporal product density.

Variance

The variance of the second-order product density estimator (3.14) can be obtained by the direct application of the extended Campbell's Theorem (Illian et al.; 2008; Chiu et al.; 2013) for the spatio-temporal case. In particular, we have

$$\mathbb{E}\left[\left(\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t)\right)^2\right] = \frac{(c(r,t))^2}{16}\left[4E_1(B) + 2E_2(B) + E_3(B)\right] \tag{3.17}$$

with

$$c(r,t) = \frac{1}{\pi r\gamma_W(r)\gamma_T(t)}$$

86

and

$$E_1(B) = \int\limits_{B^{\otimes 3}} \kappa_{1\epsilon}(\|\mathbf{x} - \mathbf{y}\| - r)\kappa_{1\epsilon}(\|\mathbf{x} - \mathbf{z}\| - r)\kappa_{2\delta}(|\xi - \eta| - t)\kappa_{2\delta}(|\xi - \zeta| - t)$$
$$\times \rho^{(3)}((\|\mathbf{x} - \mathbf{y}\|, |\xi - \eta|), (\|\mathbf{x} - \mathbf{z}\|, |\xi - \zeta|))\mathrm{d}(\mathbf{x}, \xi)\mathrm{d}(\mathbf{y}, \eta)\mathrm{d}(\mathbf{z}, \zeta),$$

$$E_2(B) = \int\limits_{B^{\otimes 2}} \kappa_{1\epsilon}^2(\|\mathbf{x} - \mathbf{y}\| - r)\kappa_{2\delta}^2(|\xi - \eta| - t)$$
$$\times \rho^{(2)}(\|\mathbf{x} - \mathbf{y}\|, |\xi - \eta|)\mathrm{d}(\mathbf{x}, \xi)\mathrm{d}(\mathbf{y}, \eta),$$

$$E_3(B) = \int\limits_{B^{\otimes 4}} \kappa_{1\epsilon}(\|\mathbf{x} - \mathbf{y}\| - r)\kappa_{1\epsilon}(\|\mathbf{z} - \mathbf{w}\| - r)\kappa_{2\delta}(|\xi - \eta| - t)\kappa_{2\delta}(|\zeta - \gamma| - t)$$
$$\times \rho^{(4)}((\|\mathbf{x} - \mathbf{y}\|, |\xi - \eta|), (\|\mathbf{x} - \mathbf{z}\|, |\xi - \zeta|), (\|\mathbf{x} - \mathbf{w}\|, |\xi - \eta|))$$
$$\times \mathrm{d}(\mathbf{x}, \xi)\mathrm{d}(\mathbf{y}, \eta)\mathrm{d}(\mathbf{z}, \zeta)\mathrm{d}(\mathbf{w}, \gamma).$$

Finding an expansion for the variance in terms of (ϵ, δ) will require knowledge of the form of the third and fourth-order product density function for a given point process model.

3.3.2 Expectation and variance of the product density estimator under Poisson process

Expectation

For a Poisson process with intensity ρ, the nth-order product density $\rho^{(n)}$ is equal to ρ^n, so utilizing (3.16) when $(\epsilon, \delta) \to (0, 0)$,

$$\mathbb{E}\left[\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t)\right] = \int\limits_{-r/\epsilon}^{\infty}\int\limits_{-t/\delta}^{\infty} \frac{\kappa_1(u)\kappa_2(v)\gamma_W(r+u\epsilon)\gamma_T(t+\delta v)}{r\gamma_W(r)\gamma_T(t)} \tag{3.18}$$

$$\times \rho^{(2)}(r+u\epsilon, t+\delta v)(r+u\epsilon)\mathrm{d}u\,\mathrm{d}v = \rho^2,$$

if the lower bound for the value of κ_1 and κ_2 are respectively larger than $-r/\epsilon$ and $-t/\delta$.

Variance

Considering (3.17) and unbiasedness property of the product density estimator, we have that

$$\mathbb{V}\mathrm{ar}\left[\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t)\right] = \frac{(c(r,t))^2}{16}\left[4\rho^3 S_1 + 2\rho^2 S_2\right], \tag{3.19}$$

where

$$S_1 = \int\limits_{B^{\otimes 3}} \kappa_{1\epsilon}(\|\mathbf{x}-\mathbf{y}\|-r)\kappa_{1\epsilon}(\|\mathbf{x}-\mathbf{z}\|-r)\kappa_{2\delta}(|\xi-\eta|-t)$$

$$\times \kappa_{2\delta}(|\xi-\zeta|-t)\,\mathrm{d}(\mathbf{x},\xi)\mathrm{d}(\mathbf{y},\eta)\mathrm{d}(\mathbf{z},\zeta)$$

$$= \int\limits_{B}\left\{\int\limits_{B} \kappa_{1\epsilon}(\|\mathbf{x}-\mathbf{y}\|-r)\kappa_{2\delta}(|\xi-\eta|-t)\,\mathrm{d}(\mathbf{y},\eta)\right\}^2 \mathrm{d}(\mathbf{x},\xi) = S_1^s S_1^t$$

and

$$S_2 = \int\limits_{B^{\otimes 2}} \kappa_{1\epsilon}^2(\|\mathbf{x}-\mathbf{y}\|-r)\kappa_{2\delta}^2(|\xi-\eta|-t)\mathrm{d}(\mathbf{x},\xi)\mathrm{d}(\mathbf{y},\eta) = S_2^s S_2^t.$$

Here,

$$S_1^s = \int\limits_{W}\left\{\int\limits_{W} \kappa_{1\epsilon}(\|\mathbf{x}-\mathbf{y}\|-r)\,\mathrm{d}\mathbf{y}\right\}^2 \mathrm{d}\mathbf{x},$$

$$S_1^t = \int_T \left\{ \int_T \kappa_{2\delta}(|\xi - \eta| - t)\, \mathrm{d}\eta \right\}^2 \mathrm{d}\xi,$$

$$S_2^s = \int_{W^{\otimes 2}} \kappa_{1\epsilon}^2(\|\mathbf{x} - \mathbf{y}\| - r)\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}$$

and

$$S_2^t = \int_{T^{\otimes 2}} \kappa_{2\delta}^2(|\xi - \eta| - t)\mathrm{d}\xi\mathrm{d}\eta.$$

For spatial case, using Epanechnikov kernel Stoyan et al. (1993b) have shown that

$$S_2^s = \frac{6}{5\epsilon}\left(|W|\pi r - U(W)\left(\frac{\epsilon^2}{7} + r^2\right) \right)$$

and

$$S_1^s = 4\pi^2 r^2\big(|W| - A\big) + 4(r + \epsilon)^2(\pi - 1)^2 A,$$

where $A = U(W)(r + \epsilon) - 4(r + \epsilon)^2$.

For temporal case and using uniform kernel, by the same methods as in Stoyan et al. (1993b), it is easy to show that

$$S_2^t = \frac{|T|}{\delta}$$

and

$$S_1^t = 4|T| - 8(t + \delta) + \frac{128}{3}t^2(t + \delta).$$

By combining the above formulas an approximation of the variance of the product density estimator is obtained. In practice we substitute ρ by its estimate $\hat{\rho} = \frac{N(W \times T)}{|W||T|}$.

3.4 Moments of the product density estimate: under the hypothesis of spatio-temporal separability

For non-negative Borel functions h_1 and h_2 defined on $(\mathbb{R}^2)^{\otimes n}$ and $\mathbb{R}^{\otimes n}$ respectively, we assume that

$$h((\mathbf{u}_1, s_1), \ldots, (\mathbf{u}_n, s_n)) = h_1(\mathbf{u}_1, \ldots, \mathbf{u}_n) h_2(s_1, \ldots, s_n),$$

and considering n-order spatio-temporal separability we can rewrite (3.1) as

$$\mathbb{E} \sum_{\Theta_n}^{\neq} h((\mathbf{u}_1, s_1), \ldots, (\mathbf{u}_n, s_n)) =$$

$$\mathbb{E} \sum_{\mathbf{u}_1, \ldots, \mathbf{u}_n \in X_{\text{space}}}^{\neq} f_1(\mathbf{u}_1, \ldots, \mathbf{u}_n) \mathbb{E} \sum_{s_1, \ldots, s_n \in X_{\text{time}}}^{\neq} f_2(s_1, \ldots, s_n),$$

$$= \int_{W^{\otimes n}} h_1(\mathbf{u}_1, \ldots, \mathbf{u}_n) \bar{\rho}_1^{(n)}(\mathbf{u}_1, \ldots, \mathbf{u}_n) \prod_{i=1}^{n} d\mathbf{u}_i$$

$$\times \int_{T^{\otimes n}} h_2(s_1, \ldots, s_n) \bar{\rho}_2^{(n)}(s_1, \ldots, s_n) \prod_{i=1}^{n} ds_i \quad (3.20)$$

where $f_1(\mathbf{u}_1, \ldots, \mathbf{u}_n) = h_1(\mathbf{u}_1, \ldots, \mathbf{u}_n)/(\int_{W^{\otimes n}} \bar{\rho}_1^{(n)}(\mathbf{u}_1, \ldots, \mathbf{u}_n) \prod_{i=1}^{n} d\mathbf{u}_i)$ and $f_2(s_1, \ldots, s_n) = h_2(s_1, \ldots, s_n)/(\int_{T^{\otimes n}} \bar{\rho}_2^{(n)}(s_1, \ldots, s_n) \prod_{i=1}^{n} ds_i)$.

3.4.1 Expectation

Combining (3.20) and (3.14) for $n = 2$ we have,

$$\mathbb{E}\left[\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t)\right] = \frac{(n-2)!}{n!}$$

$$\mathbb{E}\sum_{\mathbf{u}_i,\mathbf{u}_j\in X_{\text{space}}}^{\neq}\frac{\kappa_{1\epsilon}(\|\mathbf{u}_i-\mathbf{u}_j\|-r)}{2\pi\gamma_W(r)r}\mathbb{E}\sum_{s_i,s_j\in X_{\text{time}}}^{\neq}\frac{\kappa_{2\delta}(|s_i-s_j|-t)}{2\gamma_T(t)}$$

$$=\int_{-r/\epsilon}^{\infty}\int_{-t/\delta}^{\infty}\frac{\kappa_1(u)\kappa_2(v)\gamma_W(r+\epsilon u)\gamma_T(t+\delta v)}{r\gamma_W(r)\gamma_T(t)}\rho^{(2)}(r+\epsilon u,t+\delta v)(r+\epsilon u)\mathrm{d}u\,\mathrm{d}v$$

Note that under separability we obtain the same expression (3.16) as in the general case.

3.4.2 Second-order moment

Under the same assumptions as in the above case for $n=3,4$ in (3.20), and using (3.14) the second-order moment of the product density estimate the under separability hypothesis is given by

$$\mathbb{E}\left[\left(\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t)\right)^2\right]=\frac{(c(r,t))^2}{16}\left[2E_1(B)+4E_2(B)+E_3(B)\right].$$

For this case we also obtain expression (3.17) as in the general case. Thus all results and properties for the general case are also satisfied in the separable case.

3.5 Simulation study

The spatio-temporal second-order product density function is of interest as can be used to discriminate amongst several spatio-temporal point structures. For example, for a Poisson process $\mathbb{E}[\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t)]=\rho^2$ as we have shown previously in (3.18). Values of the spatio-temporal surface of the second-order product density function larger than the values of the plane ρ^2, indicate that the interevent distances around (r,t) are relatively more frequent compared to those in a Poisson process, which is typical of a

cluster process, and conversely, values of the spatio-temporal surface of $\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t)$ smaller than the values of the plane ρ^2 indicate that the corresponding distances are rare and this is typical of an inhibition process. The second-order product density function can take all values between zero and infinity.

We conducted a simulation experiment to analyse the behaviour of our estimator of the second-order spatio-temporal product density function under random Poisson structures. In addition, as we have developed close expressions for the variance under the Poisson case, we use them to generate the corresponding confidence surfaces. We considered the volume $W \times T = [0,10]^2 \times [0,10]$ and simulated spatio-temporal point patterns with a varying expected number of points $\mathbb{E}[N(W \times T)] = n = 100, 200, 300$. We considered $N_{\text{sim}} = 100$ repetitions per pattern and scenario. The work has been implemented in R, and has used the **stpp** package Gabriel et al. (2012). We used a fine grid for each spatial and temporal distances u and v spanning the sequence starting from $u > \epsilon > 0$ to 2.50 and $v > \delta > 0$ to 2.50 with small increments of distances. In the spatial case, Fiksel (1988a) suggested the use of the Epanechnikov kernel with bandwidth parameter $\epsilon = 0.1\sqrt{5/\rho}$. In practice, we use the **dpik** function in **kernsmooth** package to obtain the bandwidth (Wand and Ripley; 2013) based on the distances between the spatial locations of the process. For the temporal case the uniform kernel is used, where again we calculate the bandwidth δ using the **dpik** function based on the time lag between the temporal instants of the process. Note that the second-order product density function was evaluated for any scenario and repetition over the same spatio-temporal grid.

Table 1 shows some descriptive measures of the second-order spatio-temporal product density kernel estimator for homogeneous Poisson processes, under several simulated scenarios and different expected number of points. The homogeneous Poisson processes are simulated using the **rpp** function with constant intensity. The spatial and temporal bandwidths are estimated for each one of the one hundred repetitions. Table 1 displays the average optimal bandwidths for each sample size (named "Est" in Table 1).

We also show the behaviour of the second-order product density under two other fixed bandwidth values designed to overestimate and underestimate the optimal values (namd "Fix" in Table 1). From all possible grid cells, in Table 1 we have only shown the descriptive measures for some particular values of (r, t), for comparison purposes.

Table 1 also shows the theoretical second-order product density under a Poisson case ($\rho^{(2)} = \widehat{\rho^2}$), together with the estimated average surface $(\widehat{\rho^{(2)}})$, $(Q_{5\%}(\widehat{\rho^{(2)}}))$ and $(Q_{95\%}(\widehat{\rho^{(2)}}))$ are the 5% and 95% sample quantile values. In terms of variances, we present the average approximate theoretical standard deviation surface $(\sigma(\widehat{\rho^{(2)}}))$ together with the average empirical standard deviation surface values $(\widehat{\sigma}(\widehat{\rho^{(2)}}))$.

Table 1 shows the results for only three selected cells over the fine grid of spatial and temporal distances to save space. The estimated product density function over the whole grid is depicted in Figures 3.2, 3.3 and 3.4. We note that in general the difference between the estimated product density and the theoretical one is smaller when using the estimated bandwidth using dpik, compared with those cases where we use some other fixed values for the bandwidth. In addition, the variances coming from our theoretical developed expression are in the same order of the empirical variance for the selected cells, and even lower for many other cells.

## 3.6 Invasive Meningococcal Disease (IMD): Second-order analysis

Meyer et al. (2012) quantified the transmission dynamics of the two most common meningococcal antigenic sequence types observed in Germany between 2002 and 2008. They modelled the conditional intensity function by a superposition of additive and multiplicative components in space and time. According to them, the Invasive Meningococcal Disease (IMD) is a known life-threatening human disease involving meningitis (50% of cases), septicemia (5% to 20%) and/or pneumonia (5% to 15%) caused by the infection with the bacterium Neisseria meningitidis, also termed meningococ-

| n	Type	ϵ	δ	r	t	$\rho^{(2)} = \widehat{\rho^2}$	$Q_{5\%}(\widehat{\rho^{(2)}})$	$\widehat{\rho^{(2)}}$	$Q_{95\%}(\widehat{\rho^{(2)}})$	$\sigma(\widehat{\rho^{(2)}})$	$\widehat{\sigma}(\widehat{\rho^{(2)}})$
100	**Fix.**	0.70	0.15	1.1610	0.6192	0.0099	0.0097	0.0103	0.0160	0.0039	0.0030
				1.6631	1.3245	0.0099	0.0097	0.0104	0.0162	0.0063	0.0034
				2.1653	2.0298	0.0099	0.0097	0.0100	0.0157	0.0115	0.0032
	Est.	0.9936	0.3841	1.1610	0.6192	0.0099	0.0101	0.0102	0.0145	0.0038	0.0025
				1.6631	1.3245	0.0099	0.0099	0.0101	0.0149	0.0069	0.0027
				2.1653	2.0298	0.0099	0.0099	0.0099	0.0154	0.0128	0.0028
	Fix.	1.20	0.50	1.1610	0.6192	0.0099	0.0101	0.0100	0.0140	0.004	0.0023
				1.6631	1.3245	0.0099	0.0100	0.0100	0.0144	0.0075	0.0026
				2.1653	2.0298	0.0099	0.0098	0.0099	0.0151	0.0138	0.0027
200	**Fix.**	0.48	0.07	1.1610	0.6192	0.0398	0.0405	0.0407	0.0576	0.0113	0.0107
				1.6631	1.3245	0.0398	0.0414	0.0413	0.0584	0.0165	0.0105
				2.1653	2.0298	0.0398	0.0398	0.0408	0.0582	0.0293	0.0100
	Est.	0.7383	0.2466	1.1610	0.6192	0.0398	0.0399	0.0403	0.0525	0.0093	0.0067
				1.6631	1.3245	0.0398	0.0405	0.0402	0.0518	0.0167	0.007
				2.1653	2.0298	0.0398	0.0396	0.0402	0.0510	0.0313	0.0069
	Fix.	0.90	0.36	1.1610	0.6192	0.0398	0.0395	0.0398	0.0509	0.0096	0.0062
				1.6631	1.3245	0.0398	0.0392	0.0399	0.0508	0.0178	0.0066
				2.1653	2.0298	0.0398	0.0396	0.0400	0.0497	0.0333	0.0064
300	**Fix.**	0.42	0.03	1.1610	0.6192	0.0897	0.0923	0.0935	0.1305	0.0248	0.0238
				1.6631	1.3245	0.0897	0.0914	0.0944	0.1394	0.0326	0.0236
				2.1653	2.0298	0.0897	0.0946	0.0948	0.1293	0.0546	0.0196
	Est.	0.6093	0.1862	1.1610	0.6192	0.0897	0.0934	0.0933	0.1144	0.0166	0.0135
				1.6631	1.3245	0.0897	0.0941	0.0947	0.1187	0.0295	0.0157
				2.1653	2.0298	0.0897	0.0942	0.0953	0.1219	0.0556	0.0144
	Fix.	0.81	0.34	1.1610	0.6192	0.0897	0.0923	0.0926	0.1124	0.0171	0.0122
				1.6631	1.3245	0.0897	0.0931	0.0936	0.1160	0.0319	0.0137
				2.1653	2.0298	0.0897	0.0929	0.0944	0.1161	0.0600	0.0135

Table 3.1: *Descriptive measures of the estimation of the second-order spatio-temporal product density under the Poisson case.*
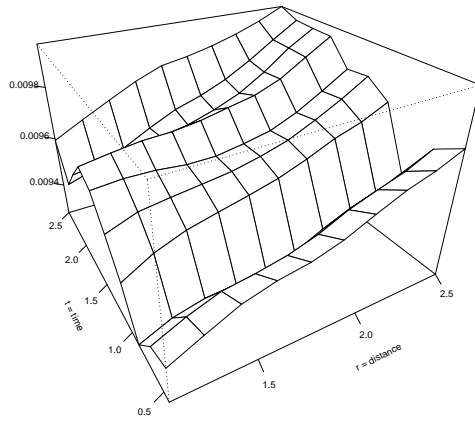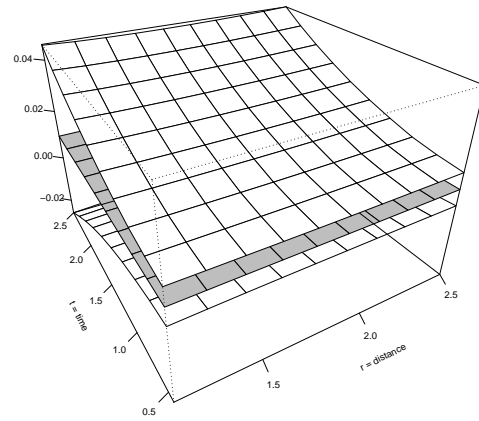
(a) Expected average surface.
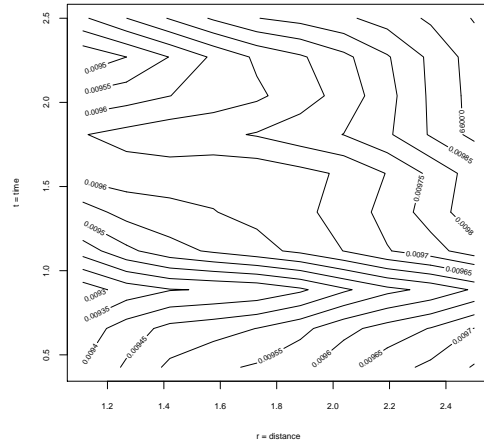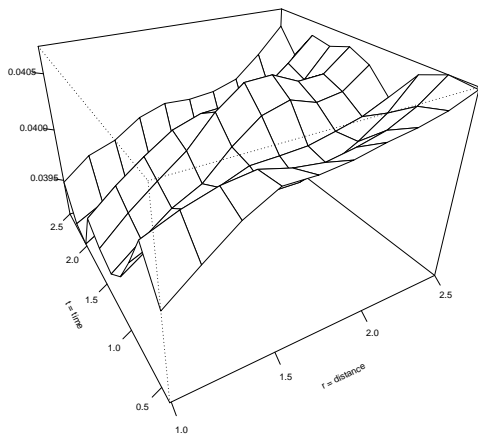
(b) Expected average contours.
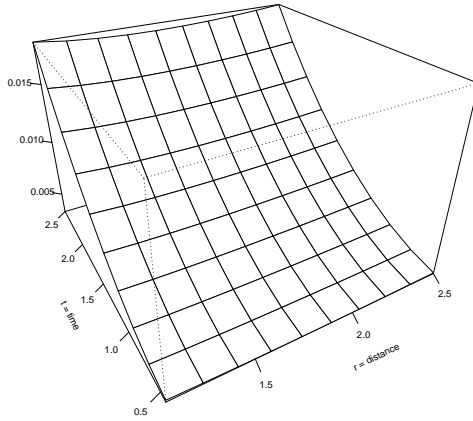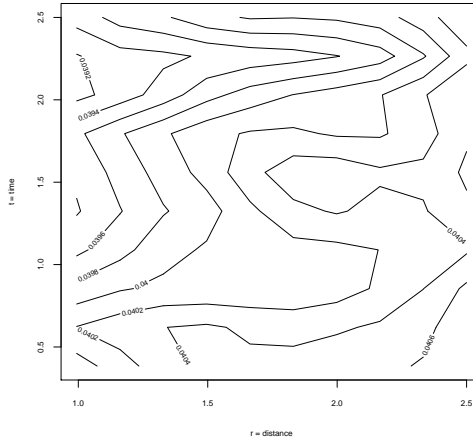

(c) Average standard deviation surface.

(d) Average confidence surfaces.

Figure 3.2: *Statistical properties of the scond-order spatio-temporal product density kernel estimated under Poisson point patterns with expected number of points $n = 100$, $\widehat{\rho^2} = 0.0099$, $\epsilon = 0.9936$ and $\delta = 0.3841$.*

cus. Meningococci can be transmitted airborne (by coughing or sneezing) or by other mucous secretions from infected humans, where they colonise the nasopharynx, their only natural reservoir. The risk of contracting IMD is much higher inside the household of an infected person. Nevertheless, secondary cases have become rare due to effective antimicrobial chemopro phylaxis of household members and anyone exposed to an infected patient with oral secretions. The risk of secondary infections is highest during the first few days. Meyer et al. (2012) claim that most meningococci are com-
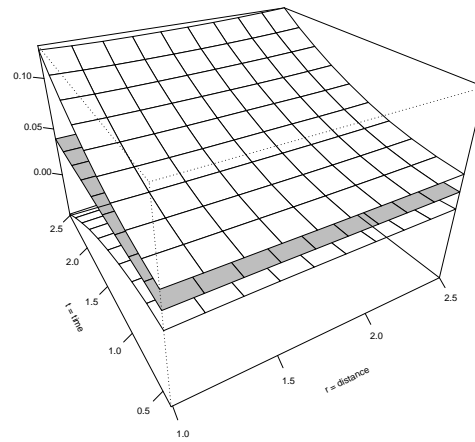
(a) Expected average surface.

(b) Expected average contours.
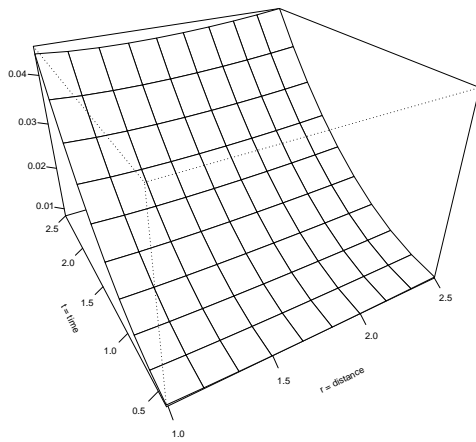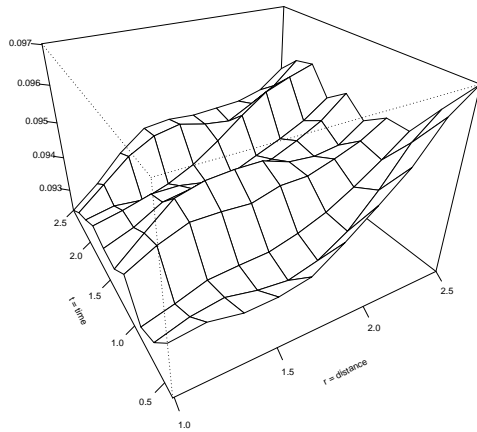

(c) Average standard deviation surface.

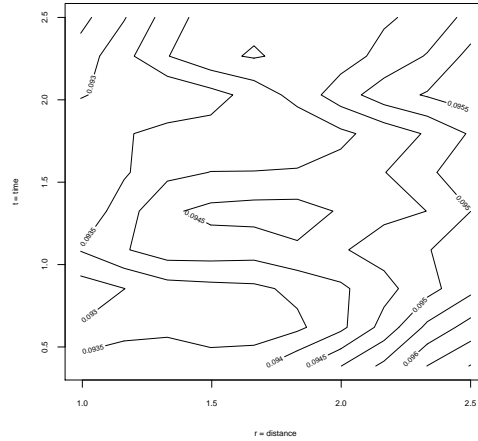(d) Average confidence surfaces.

Figure 3.3: *Statistical properties of the scond-order spatio-temporal product density kernel estimated under Poisson point patterns with expected number of points $n = 200$, $\widehat{\rho^2} = 0.0398$, $\epsilon = 0.7383$ and $\delta = 0.2466$.*

mensal in humans, but only a few isolates are virulent and cause invasive disease.

The area of Germany is 357603 km^2 with a perimeter of 6146 km. The IMD dataset consists of the spatio-temporal reports of 636 cases of IMD caused by two specific meningococcal finetypes in which the times are given by 2569 days over the 7-year period, so the temporal region is defined as $T = [0, 2569]$. Figure 3.5 shows the estimated spatial intensity
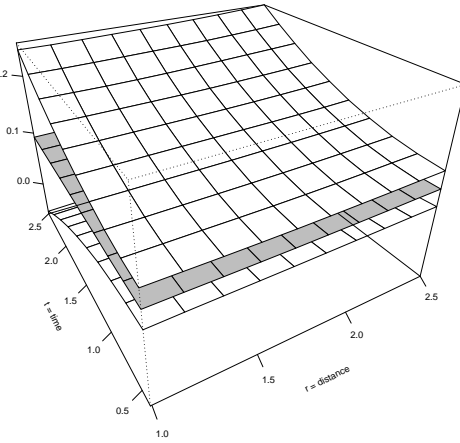
96

(a) Expected average surface.

(b) Expected average contours.
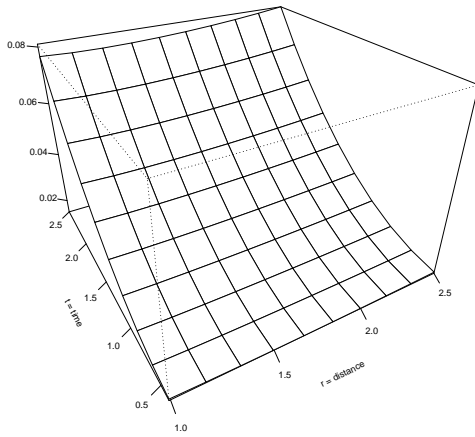
(c) Average standard deviation surface.

(d) Envelope surface surfaces.

Figure 3.4: *Statistical properties of the scond-order spatio-temporal product density kernel estimated under Poisson point patterns with expected number of points $n = 300$, $\widehat{\rho^2} = 0.0897$, $\epsilon = 0.06093$ and $\delta = 0.01862$.*

(left panel) and estimated temporal intensity (right panel). In the purely spatial case, this figure shows clearly the inhomogeneity condition of IMD, with a notorious high intensity of points per km^2 in the western border of Germany, and some lower intensity (but noticeable concentrations) near the north-eastern and southern borders. Figure 3.6 shows the surface of the estimated second-order product density using $\epsilon = 13.9686$ km and $\delta = 28$ days. This figure shows large values for small spatial and temporal
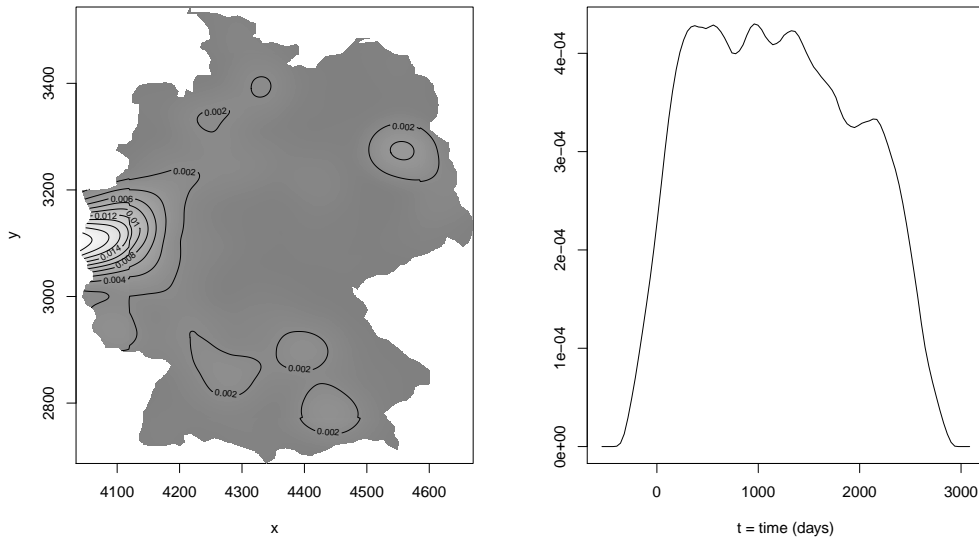
Figure 3.5: *Estimated spatial (left panel) and temporal (right panel) intensities for the IMD dataset.*

distances, which is a typical behaviour of a cluster spatio-temporal point pattern. However, the spatial aggregation decreases with increasing spatial distances, while the temporal aggregation is kept throughout most of the temporal range, as clearly shown in the right panel of Figure 3.5. This result is a consequence of many reports of IMD occurring close in space and time, and thus for short temporal periods it is quite likely that at least two reports of IMD occur close enough of each other. Additionally, the spatial aggregation shows the same behaviour even during periods of time sufficiently large. One way to emphasise this clustering behaviour is to compare the theoretical value of the second-order product density function for a Poisson point pattern with equal expected number of points than IMD ($\hat{\rho^2} = 4.840609 \times 10^{-13}$), with the maximum value of the empirical surface in Figure 3.6, which is around 1.2×10^{12}. This result is clearly expected after visual inspection of Figure 3.6, and goes in the line found by Meyer et al. (2012)).

The left panel of Figure 3.7 shows the 95%-envelope surfaces obtained from 39 simulations of a spatio-temporal Poisson point pattern, see Møller and Ghorbani (2010) and Møller and Waagepetersen (2004)), together with

98
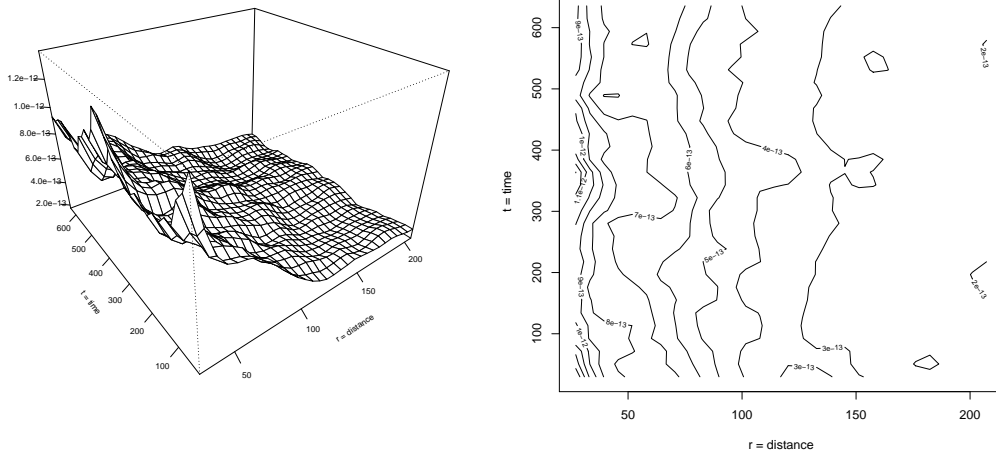
Figure 3.6: *Second-order product density kernel estimator for the spatio-temporal IMD dataset with $\hat{\rho}^2 = 4.893802 \times 10^{-13}$, $\epsilon = 13.9686$ km, and $\delta = 28$ days.*
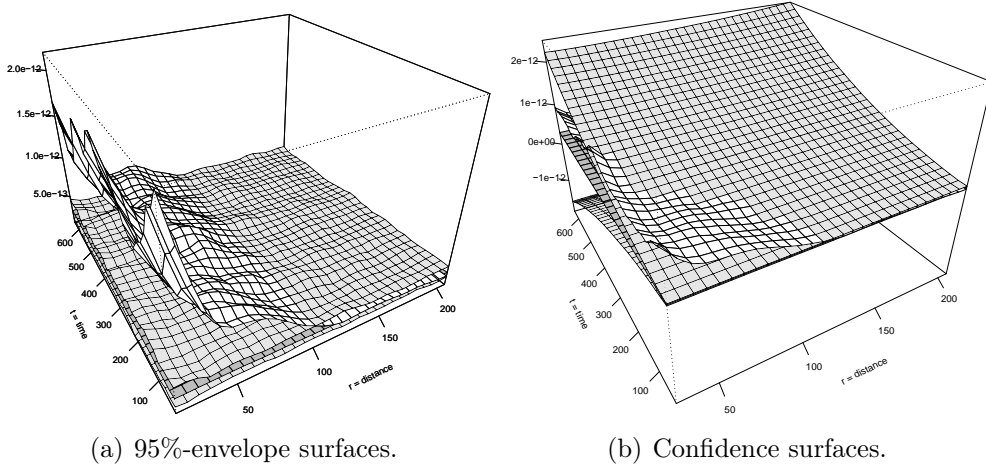
(a) 95%-envelope surfaces.

(b) Confidence surfaces.

Figure 3.7: *3.7(a) 95%-envelope surfaces obtained from 39 simulations of a spatio-temporal Poisson point pattern; 3.7(b) confidence interval under a Poisson pattern based on the estimated $\widehat{\rho^{(2)}} \pm$ two standard deviations calculated using the close form of the variance in Section 3.3.1 In both cases, we superimpose the empirical product density for the IMD data estimated using $\epsilon = 13.9686$ km and $\delta = 28$ days.*

the empirical second-order product density. This figure shows how the empirical surface of the second-order product density function for the IMD dataset is larger than the upper 95%-envelope for small spatial and temporal distances.

The right panel of Figure 3.7 shows the confidence interval under a Poisson pattern based on the estimated $\widehat{\rho^{(2)}}$ ± two standard deviations calculated using the close form of the variance in Section 3.2. We also superimpose the empirical product density for the IMD data. Again, the empirical density goes out the upper confidence surface. These two figures reveal that IMD has a contagious behaviour in their immediate spatio-temporal neighborhoods. These are solid arguments to reject the hypothesis of complete randomness in favour of a clustering structure.

3.7   Discussion

The spatio-temporal inhomogeneous second-order product density function describes second-order characteristics of point processes. It is useful to analyse the spatio-temporal structure of the underlying point process, and thus provides a natural starting point for the analysis of spatio-temporal point process data. It can be considered an exploratory tool, for testing spatio-temporal clustering or spatio-temporal interaction.

We have proposed a non-parametric edge-corrected kernel estimate of the product density under the second-order intensity-reweighted stationary hypothesis. The expectation and variance of the estimator are obtained, and close expressions are derived under the Poisson case. First- and second-order spatio-temporal separability has also been considered and discussed. It is known (see Gabriel (2013)) that the performance of the pair correlation function and K-function can be severely altered by the intensity estimate. This can be explained by over-parametrisation or over-fitting in the case of a parametric estimation of the intensity function, or by the incapacity of distinguish first- and second-order effects from a single realisation of the point process in the case of a kernel-based estimation. This is in any case a kind of handicap and weakness in using these characteristics. We

postulate the use of the second-order product density as it provides the same amount of information, but with the added value that there is no need to estimate the intensity function.

We have provided sufficiently statistical grounds in favour of using this second-order tool in the practical analysis of spatio-temporal point patterns. However, we have based under developments on the hypothesis of second-order intensity-reweighted stationarity. The statistical properties of the spatio-temporal product density under general non-stationarity conditions or anisotropic structures remains an open problem.

Second-order analysis of anisotropic spatio-temporal point process data

Carles Comas[a], Francisco J. Rodríguez-Cortés[b] and
Jorge Mateu[b],[1]

[a] Department of Mathematics, Universitat de Lleida, Lleida, Spain
[b] Department of Mathematics, Universitat Jaume I, Castellón, Spain

Abstract

Second-order orientation methods provide a natural tool for the analysis of spatial point process data. In this paper we extend to the spatio-temporal setting the spatial point pair orientation distribution function. The new spatio-temporal orientation distribution function is used to detect spatio-temporal anisotropic configurations. An edge-corrected estimator is defined and illustrated through a simulation study. We apply the resulting estimator to data on the spatio-temporal distribution of fire ignition events caused by humans in a square area of 30×30 km^2 during four years. Our results confirm that our approach is able to detect directional components at distinct spatio-temporal scales.

---

[1]All authors contributed equally in this work.

102

4.1 Introduction

Spatial point process models are useful tools to model irregularly scattered point patterns that are frequently encountered in biological, ecological, and epidemiological studies. A spatial point pattern is a set of points $\{\mathbf{u}_i \in W : i = 1, \ldots, n\}$ in some planar region W. Very often, W is a sampling window within a much larger region and it is reasonable to regard the point pattern as a partial realization of a stochastic planar point process, the events consisting of all points of the process which lie within W. Recent textbooks related to the topic of analysis and modeling of point processes include Stoyan et al. (1995), Diggle; 2013, Baddeley et al. (2006), Illian et al. (2008), or Gelfand et al. (2010).

Today, much attention is paid to spatio-temporal point processes, where each point represents the location and time of an event, and thus we have data of the form $(\mathbf{u}_i, s_i) \in W \times T$, $i = 1, ..., n$, with $T \in \mathbb{R}$ denoting the temporal interval. There has been a lot of recent work on spatio-temporal models, and a variety of ad-hoc approaches have been suggested (Gelfand et al. (2010); Diggle (2013)). Processes that are both spatially and temporally discrete are more naturally considered as binary-valued random fields. Processes that are temporally discrete with only a small number of distinct event-times can be considered initially as multivariate point processes. Conversely, spatially discrete processes with only a small number of distinct event-locations can be considered as multivariate temporal point processes, but with a spatial interpretation to the component processes. The other more common end, and the one considered here, is considering processes that are temporally continuous and either spatially continuous or spatially discrete on a sufficiently large support to justify formulating explicitly second-order spatio-temporal tools for the data.

A point process is stationary and isotropic if its statistical properties do not change under translation and rotation, respectively. Informally, *stationarity* implies that one can estimate properties of the process from a single realization on $W \times T$, by exploiting the fact that these properties are the same in different, but geometrically similar, subregions of $W \times T$; *isotropy* means that there are no directional effects.

The assumption of isotropy is often made in practice due to a simpler interpretation, ease of analysis, and also to increase the power of statistical analyses. However, isotropy is many times hard to find in real applications. Many point processes are indeed anisotropic. There are many varied forms of anisotropy. Orientation analysis is the quantification of the degree of anisotropy in the case of anisotropic point patterns and the detection of inner orientations in case of isotropy (Ohser and Stoyan (1981); Stoyan and Beneš (1991); Mateu (2000); Redenbach et al. (2009)). Typical examples of oriented point patterns are patterns in which the points lie randomly in parallel strips of random or constant breadth (anisotropic case) or on an isotropic system of random fibres (inner orientation). Anisotropy is the converse of isotropy but it has many different aspects. For example, the anisotropy (directionality) is a characteristic property of images. Anisotropy may be the result of the process by which the imaged object might have been formed. Thus, on numerous occasions anisotropy reflects properties and determines the behavior of the textured objects. The importance of anisotropy in visual perception and object characterization inspired a range of studies for anisotropy analysis (Kovalev and Bondar (1997)). Anisotropy can be present when the spatial point patterns contain points placed roughly on line segments. See details in Møller and Rasmussen (2012) who consider a particular class of point processes whose realizations contain such linear structures.

Ohser and Stoyan (1981) presented a method for the second-order analysis of anisotropic point processes. They provided expressions for the anisotropic spatial K-function, and used this function to define an orientation distribution function. This represented a simple and intuitive approach to obtain information on the anisotropic properties of the spa-

tial distribution of a point pattern. Later, Rosenberg (2004) proposed methods to assess isotropy (and to consequently detect anisotropy) for spatial point processes. Rosenberg (2004) analyzed the second-order properties by anisotropic techniques to study the directional relationship among the observed points, and proposed a geometric anisotropic pair-correlation function. This approach, however, is limited to certain classes of models exhibiting close forms for certain second-order statistics. Guan et al. (2004, 2006) proposed a formal nonparametric approach to test for isotropy based on the asymptotic joint normality of the sample second-order intensity function. Alternative methods based on two-dimensional spectral analysis were proposed by Mugglestone and Renshaw (1998) to calculate objective estimates of the orientation and frequency of geological lineations from digitized images obtained from aerial photographs of glaciated terrain in northern Canada. The complications inherent in spectral analysis (particularly for more than one dimension) appear to have discouraged applied statisticians and ecologists from making use of these methods. Wavelet analysis has succeeded in a variety of applications and held promise in the area of spatial pattern analysis (e.g.Donoho (1993); Gao and Li (1993); Grenfell et al. (2001)). However, wavelet analysis has only been involved in several works for detection of spatial patterns (e.g. Harper and Macdonald (2001); Perry et al. (2002)), but we can not find connections with spatio-temporal data. In general there is not a large treatment of anisotropy for spatial planar point patterns, and there is currently only one such approach for three-dimensional point patterns (Redenbach et al. (2009)).

In this paper, we consider data in the form of a realization of a spatio-temporal point process within a finite spatio-temporal region. Many spatial processes of scientific interest also have a temporal component that may need to be considered when modeling the underlying phenomenon (e.g., distribution of cases for a disease or assessment of risk of air pollution). Spatio-temporal point processes, rather than purely spatial point processes, must then be considered as potential models. There is an extensive literature on the analysis of point process data in time (e.g., Cox and Isham (1980); Daley and Vere-Jones (2008)) and in space (e.g., Cressie

(1993); Diggle (2013); Møller and Waagepetersen (2004)). Generic methods for the analysis of spatio-temporal point processes are less well established; see for example Diggle (2006), Gabriel and Diggle (2009), Møller and Ghorbani (2012), Section 6.6 of Cressie and Wikle (2011), and Gabriel (2013).

Any current approach to the analysis of spatio-temporal point patterns assumes that the process is isotropy at all times, sometimes considering it is stationary (with a constant trend), and others non-stationary (and then controlling for the inhomogeneity degree). We are interested in analyzing anisotropic properties in form of orientations in spatio-temporal point patterns. These patterns are more realistic when modeling forest fires, as shown in this paper, or in general when modeling a wide variety of environmental patterns.

Our aim is to extend, to the spatio-temporal setting, the orientation analysis provided by Ohser and Stoyan (1981) in their early contribution for spatial planar patterns. The new spatio-temporal orientation distribution function is used to detect spatio-temporal anisotropic configurations. We define an edge-corrected estimator, and its practical behavior is illustrated through a simulation study, and an application to fire events. The plan of the paper is the following. In Section 4.2 we present a brief methodological setup of spatio-temporal point patterns, with an emphasis on an anisotropic spatio-temporal second-order characteristic. In Section 4.3 we introduce an orientation analysis for such anisotropic spatio-temporal patterns. A simulation study is presented in Section 4.4. Finally, a case study is considered in Section 4.5 to illustrate the use of our new approach. The paper ends with a summary section where we briefly provide some conclusions.

4.2   Methodology

We are concerned with the analysis of data of the form $(\mathbf{u}_i, s_i) \in W \times T$, $i = 1, ..., n$ with $W \times T$ a subset of $\mathbb{R}^2 \times \mathbb{R}$. Each $\mathbf{u}_i$ denotes the location and s_i the corresponding time of occurrence of an event of interest. We assume

that the data form a complete record of all events which occur within a pre-specified spatial region W and time-interval T. We call a dataset of this kind a spatio-temporal point pattern, and the underlying stochastic model for the data a spatio-temporal point process. Many spatial processes of scientific interest are of this type.

A spatio-temporal point process is second-order intensity reweighted stationary (SOIRS) and isotropic if its intensity function is bounded away from zero and its pair-correlation function depends only on the spatio-temporal difference vector. Second-order intensity reweighted stationarity is defined for purely spatial point processes in Baddeley et al. (2000). Gabriel and Diggle (2009) provide the straightforward extension to the spatio-temporal case. Diggle et al. (1995) introduced in an heuristic form first and second-order spatio-temporal properties, such as the spatio-temporal intensity function and the spatio-temporal K-function. Gabriel and Diggle (2009), for a SOIRS spatio-temporal point process, provide an improved definition of the inhomogeneous spatio-temporal K-function. Møller and Ghorbani (2012) define the spatio-temporal inhomogeneous K-function in a more natural form, this definition involves the spatio-temporal pair correlation function g_0 which, according to Gabriel and Diggle (2009), can be informally interpreted as the standardized probability density that an event occurs in each of two small volumes. Note further that for a spatio-temporal Poisson process, $g_0 = 1$ and $K(r,t) = 2\pi r^2 t$. In Møller and Ghorbani (2012), Ghorbani (2013), and Gabriel (2013) alternative approximately unbiased non-parametric estimators for both the spatio-temporal inhomogeneous K-function and the intensity function are given.

Throughout this paper, we basically follow the notation and setup introduced in Møller and Ghorbani (2012) for spatio-temporal point processes. We consider a spatio-temporal point process with no multiple points as a random countable subset X of $\mathbb{R}^2 \times \mathbb{R}$, where a point $(\mathbf{u}, s) \in X$ corresponds to an event at $\mathbf{u} \in \mathbb{R}^2$ occurring at time $s \in \mathbb{R}$. In fact, we observe n events $\{(\mathbf{u}_i, s_i)\}$ of X within a bounded spatio-temporal region $W \times T \subset \mathbb{R}^2 \times \mathbb{R}$, with area $|W| > 0$, and length $|T| > 0$. For a further formal definition of a point process, see e.g. Illian et al. (2008). In our

context, $N(W \times T)$ is the number of the events of the process falling inside the bounded spatio-temporal window $W \times T$.

In the case of anisotropic planar point processes, Ohser and Stoyan (1981) defined a reduced second moment measure, and provided an estimator for the orientation analysis when the intensity is known. We consider here a similar approach but assuming SOIRS and anisotropic spatio-temporal point processes. An intuitive way of thinking about the nature of the spatio-temporal anisotropic K-function is that this function should be proportional to the mean number of points in a cylindrical sector with spatial distance r, angle φ and time lag t, centered at an arbitrary point of the spatio-temporal point process X. Then

$$K(r,t,\varphi) = \int \sum_{(\mathbf{u},s),(\mathbf{v},l) \in X}^{\neq} \mathbf{1}\left[\|\mathbf{u}\| \leq r, |\theta| \leq \varphi, |s| \leq t\right] g_0(\mathbf{u},s)\mathrm{d}(\mathbf{u},s),$$

(4.1)

wiht $r > 0$, $0 \leq \varphi \leq \pi$, $t > 0$, here $\sum^{\neq}$ means that the summation goes over all pairs $((\mathbf{u},s),(\mathbf{v},l))$ with $(\mathbf{u},s) \neq (\mathbf{v},l)$, $\mathbf{1}[\cdot]$ denotes the indicator function and $g_0(\mathbf{u},s)$ is the spatio-temporal pair-correlation function, where we set $a/0 = 0$ for $a \geq 0$, and θ denotes the least angle between the directed line from two points of the process and the x-axis. The pair (r,φ) denotes the point with polar coordinates r and φ. Note that this definition slightly differs from that of $K(r,t)$ even in the SOIRS stationary and isotropic case, see Ghorbani (2013). The reduced second moment measure K studied in Møller and Toftaker (2012), is given by $K(r,t) = K(r,t,2\pi) = 2K(r,t,\pi)$.

For a SOIRS anisotropic and stationary spatio-temporal point process, an approximately non-parametric edge-corrected estimate of the K-function (4.1) is given by

$$\widehat{K}(r,t,\varphi) = \frac{1}{|W||T|} \sum_{i=1}^{n} \sum_{i \neq j} \frac{\mathbf{1}[\|\mathbf{u}_i - \mathbf{u}_j\| \leq r, |\theta(\mathbf{u}_i,\mathbf{u}_j)| \leq \varphi, |s_i - s_j| \leq t]}{\widehat{\rho}^2 w^2(\mathbf{u}_i,\mathbf{u}_j)w^1(s_i,s_j)}$$

(4.2)

where $w_1(\mathbf{u}_i,\mathbf{u}_j)$ is the isotropic spatial edge-correction factor as proposed by Ripley (1976, 1988). We note that in the literature of spatial point

processes (see, for example, Stoyan and Stoyan (1994); Illian et al. (2008)) isotropic edge-correction factors have been used even for the anisotropic estimators in the spatial case. We follow this procedure, and make use of an isotropic spatial edge-correction factor in (4.2). The default time interval T is usually taken to be the smallest time interval containing all event times, i.e. $T = t_{\max} - t_{\min} = \max\{s_i\} - \min\{s_i\}$ whit $i = 1, \ldots, n$. For a pair of temporal observed events $s_i, s_j \in T$ where $s_j > s_i$, $w_2(s_i, s_j)$ is the temporal edge-correction factor which is equal to one if both ends of the interval of length $2|s_i - s_j|$ and center s_i lie within T, and two otherwise (Diggle et al. (1995)). An unbiased estimator for ρ^2 is $n(n-1)/(|W||T|)^2$, see Ghorbani (2013).

4.3   Orientation analysis for anisotropic spatio-temporal point processes

Based on $K(r, t, \varphi)$, we can detect predominant directions in spatio-temporal point patterns extending to the spatio-temporal domain the orientation analysis suggested by Ohser and Stoyan (1981). These authors used the directional distribution of line segments connecting point pairs of the point pattern which are a distance between r_1 and r_2 apart. The corresponding distribution function is called point pair orientation distribution (see also Stoyan and Stoyan (1994)), and is equal to the probability that a randomly chosen line segment forms an angle with the x-axis smaller than φ ($0 \leq \varphi \leq \pi$). A spatio-temporal counterpart version can be obtained via

$$O_{((r_1,t_1),(r_2,t_2))}(\varphi) = \int\limits_0^\varphi \int\limits_{t_1}^{t_2} \int\limits_{r_1}^{r_2} \mathrm{d}K(r, t, \psi) \bigg/ \int\limits_0^\pi \int\limits_{t_1}^{t_2} \int\limits_{r_1}^{r_2} \mathrm{d}K(r, t, \psi), \qquad (4.3)$$

with $r_2 > r_1 \geq 0$, $t_2 > t_1 \geq 0$. For a suitable positive value of t, $O_{((0,t),(r,t))}(\varphi)$ describes a *short-range* spatial directionality in the point pattern, $O_{((r_1,t),(r_2,t))}(\varphi)$ provides a *middle-range* spatial orientation for $r_1 < r_2$, while $O_{((r,t),(\infty,t))}(\varphi) = \lim\limits_{r_2 \to \infty} O_{((r,t),(r_2,t))}(\varphi)$ describes *long-range* spatial di-

rectionality. Further combinations of spatial and temporal intervals are not worth to be described given their lack of practical aplicability. In case of isotropy, all these distributions coincide with the uniform distribution on $[0, \pi]$.

Moreover, as (4.3) is a cumulative measure for a given angle φ, it can be useful to consider a cylindrical sector instead to better highlight the possible directional components, i.e.

$$O^*_{\phi,((r_1,t_1),(r_2,t_2))}(\varphi) = O_{((r_1,t_1),(r_2,t_2))}(\varphi+\phi) - O_{((r_1,t_1),(r_2,t_2))}(\varphi-\phi),\ 0 \leq \varphi \leq \pi,$$
$$(4.4)$$

where $\varphi > \phi > 0$ is a fixed prescribed angle interval, which provides the direction in which anisotropic effects are tested. Using (4.2) we obtain an estimator of (4.3) and (4.4), i.e. $\widehat{O}_{((r_1,t_1),(r_2,t_2))}(\varphi)$ and $\widehat{O^*}_{\phi,((r_1,t_1),(r_2,t_2))}(\varphi)$, respectively.

4.4 Simulation studies

We conducted a simulation study to illustrate the use of the spatio-temporal orientation distribution under several point configurations. For this aim, we considered several point configurations assuming distinct range of anisotropy through space and time. Here, we used spatio-temporal realizations of stationary and anisotropic Poisson point processes in $[0, 1]^2 \times [0, 1]$.

4.4.1 Generating spatio-temporal stationary and anisotropic Poisson cluster point patterns

We considered a simple birth process to generate spatio-temporal stationary Poisson cluster point processes. Let us assume that immigrants arrive randomly in time according to a Poisson process with rate ω and have uniformly distributed locations $\mathbf{u}_{\text{new}}$ on $U(0, 1)^2$. Now new arrived points are accepted or rejected in terms of the following spatio-temporal interaction mechanism. Consider $d_i = \|\mathbf{u}_i - \mathbf{u}_{\text{new}}\|$ and $z_i = s_{\text{new}} - s_i$ to be the spatial and temporal distances between a newly arrived point at time s_{new} and an

110

already established one $i = 1, \ldots n_{s_{\mathrm{new}}}$, respectively, with $n_{s_{\mathrm{new}}}$ being the total number of established points at time s_{new}. Moreover, assume $\alpha_{i,\mathrm{new}}$ to be the spatially projected angle (w.r.t. the $x-$axis) between a newly arrived immigrant and an already established point i. Now immigrants are accepted with probability

$$ P = \prod_{i=1}^{n_{s_{\mathrm{new}}}} \Pi_i \tag{4.5} $$

where

$$ \Pi_i = \begin{cases} a & \text{if } r_1 < d_i < r_2 \ \text{and} \ \ t_1 < z_i < t_2 \ \text{and} \ \ \beta + c_\beta < \alpha_{i,\mathrm{new}} < \beta - c_\beta \\ 1 & \text{otherwise,} \end{cases} $$

$$ \tag{4.6} $$

and $r_1 < r_2$ and $t_1 < t_2$ are the spatial and temporal scales of anisotropy, respectively, β is the prescribed anisotropic directional effect ($0 \le \beta \le \pi$), c_β is a tolerance angle, and $0 \le a \le 1$ is a constant that defines the strength of the anisotropic effects. The resulting process defined by (4.5) is a spatio-temporal stationary and anisotropic Poisson cluster process with constant spatio-temporal intensity $\rho = \omega$. Here, to deal with spatial edge-effects we place the spatio-temporal generation mechanism on the unit torus.

4.4.2 Monte Carlo approach to testing for spatio-temporal anisotropic effects

To test for evidence of spatio-temporal anisotropy, we followed common practice by comparing the estimator $\widehat{O^*}_{\phi,((r_1,t_1),(r_2,t_2))}(\varphi)$ with estimates obtained from simulations under a suitable null hypothesis. Here the null hypothesis is that the underlying point process cluster is a stationary and isotropic Poisson process cluster, and therefore the empirical spatio-temporal pattern is compared with a stationary and isotropic spatio-temporal Poisson process cluster with the same point intensity, based on a Monte Carlo test. We simulated 199 spatio-temporal point patterns under this null hypothesis and for each one an estimator of (4.4) is obtained. This set of functions is then compared with the resulting estimator for the em-

pirical data under analysis. Under this test, we reject the null hypothesis (spatio-temporal point isotropy) if the resulting estimator of this function lies outside the fifth-largest and/or fifth-smallest envelope values obtained from the set of simulated functions $\widehat{O^*}_{\phi,((r_1,t_1),(r_2,t_2))}(\varphi)$, with an exact significance level of $2 \times 5/(199+1) = 0.05$.

4.4.3 Simulated examples

We consider two distinct stationary and anisotropic Poisson cluster scenarios in the unit volume, namely a spatio-temporal short-range orientation structure and a spatio-temporal middle-range orientation configuration. We do so to illustrate the behavior of our new approach under distinct spatio-temporal ranges of anisotropy. Under both scenarios we take $\beta = \pi/2$ and $c_\beta = \pi/9$, as the value of the prescribed anisotropic directional effects, and a spatio-temporal intensity $\rho = 250$, to provide an enough number of points. Moreover, to generate spatio-temporal short-range orientation effects, we assume $r_1 = 0$, $r_2 = 0.2$, $t_1 = 0$ and $t_2 = 0.1$. Now points with distances in the range $0 - 0.2$ and temporal distances of less than 0.1 units are expected to have a projected spatial angle of around $\pi/2$. Also, to obtain spatio-temporal middle-range orientation structures we take $r_1 = 0.2$, $r_2 = 0.4$, $t_1 = 0$ and $t_2 = 0.1$. Now points with distances in the range $0.2 - 0.4$ and temporal distances of less than 0.1 are also expected to have a projected spatial angle of around $\pi/2$. Regarding the strength of anisotropy, we consider for the short-range scenario $a = 0$ and $a = 0.5$ to show the effect of assuming strong and weak spatio-temporal anisotropy, respectively. Notice that for $a = 0$ newly arrived points interacting with other ones in the case of spatio-temporal short-range anisotropy can only become established if their spatially projected angle with respect to any other already established point is in the interval $\beta \pm c_\beta$, whilst for $a = 0.5$ half of new arrived points will do so. For the middle-range anisotropic scenario, we only take $a = 0.5$ to avoid the quite self-evident spatio-temporal anisotropic effects obtained for $a = 0$.

Figure 4.1 shows the resulting spatio-temporal point configuration for

the short-range orientation scenario with $a = 0$, and it highlights that points ($n = 253$) are located forming several point alignments with inter-alignment distances of around 0.2, regardless of the time scale. Apparently the short-time scale of anisotropy is not visually apparent in the spatio-temporal configuration. The locations of the 253 points confirm the "column" structure for inter-event spatial distances of around 0.2 (Figure 1b). Moreover, this figure also shows the time arrival and the cumulative arrival of points (Figures 1c and d) and highlights that point establishment through time is quite constant and equals the spatio-temporal point intensity $\rho = 250$.
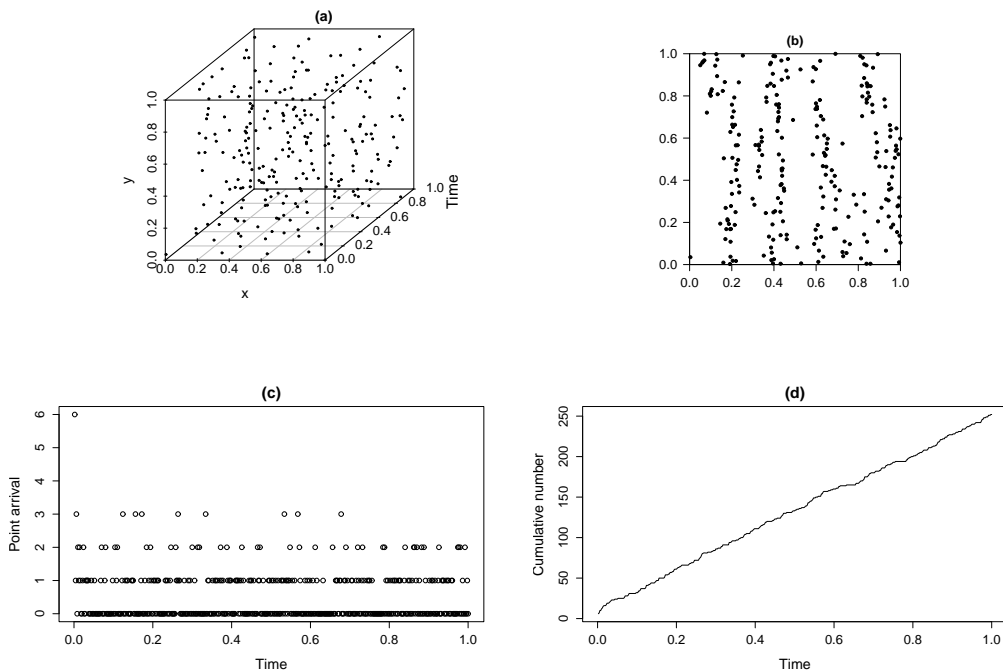

Figure 4.1: *(a) Spatio-temporal point configuration for the spatial short-range orientation scenario ($r_1 = 0$, $r_2 = 0.2$, $t_1 = 0$ and $t_2 = 0.1$), and anisotropic directional effects $\beta = \pi/2$ and $c_\beta = \pi/9$ with $a = 0$; (b) Spatial positions of points; (c) Time arrival of points, and (d) Cumulative arrival of points.*

Let us now analyze the spatio-temporal orientation structure by using our orientation distribution function $O^*_{\phi,((r_1,t_1),(r_2,t_2))}(\varphi)$ (4.4). Here this function is computed assuming the prescribed angle interval $\phi = \pi/32$ which results in 18 angle intervals. Moreover, given the potentially infi-

nite number of spatio-temporal intervals to be tested, here we explored four of them to detect and describe the scale of anisotropy. Figure 4.2a shows the empirical $\widehat{O^*}_{\phi,((0,0),(0.2,0.1))}(\varphi)$, which is the orientation distribution function for the spatio-temporal intervals used to generate this first scenario. This highlights that this empirical function lies outside the upper envelope at $\varphi = \pi/2$, thereby suggesting a spatio-temporal anisotropic structure for this angular direction. Moreover, if we explore the same spatial scale, but we increase the time interval from 0.1 to 0.3, the resulting $\widehat{O^*}_{\phi,((0,0),(0.2,0.3))}(\varphi)$ (Figure 4.2b) also suggests a spatio-temporal anisotropic structure at $\pi/2$. This result is expected because $\widehat{O^*}_{\phi,((0,0),(0.2,0.3))}(\varphi)$ is a cumulative function which also incorporates the direction effect detected for $\widehat{O^*}_{\phi,((0,0),(0.2,0.1))}(\varphi)$.

Moreover, in order to explore a temporal interval distinct to that considered to generate the first scenario, we take $t_1 = 0.1$ and $t_2 = 0.4$. Now the empirical estimator $\widehat{O^*}_{\phi,((0,0.1),(0.2,0.4))}(\varphi)$ keeps on showing orientation effects at $\pi/2$ (Figure 4.2c). This result suggests that the anisotropic structure of this first scenario is quite independent of the time scale. In fact, newly arrived points can only become established either if they do not interact with any already established points in terms of the spatio-temporal scale of anisotropy, or they do interact with other points althought their spatially projected angles with respect to these established points are in the interval $\beta \pm c_\beta$. Under this strong interaction mechanism, established points affect points that are 0.1 units apart in time, and in turn these points also affect the establishment of new immigrants. For $a = 0$, this generates a spatio-temporal structure where points tend to become established forming point alignments in the anisotropic direction with inter-alignment distances defined by the spatial scale of anisotropy. In our case, these point alignments are in the $\pi/2$ direction with inter-alignment distances of around 0.2 units. Note that this results in a configuration of points quite independent of the temporal scale. This is the reason why our orientation distribution function detects anisotropic effects for the defined spatial scale regardless of the temporal scale.

Figure 4.3 shows the resulting spatio-temporal point configuration for
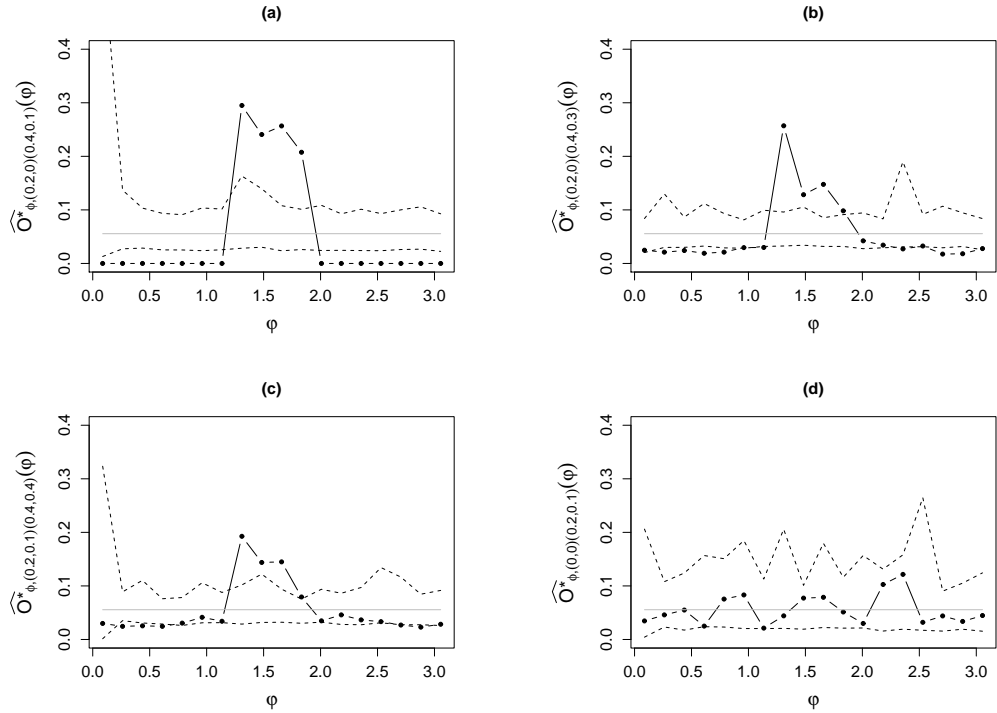
114

Figure 4.2: *Resulting estimator of the orientation function $\widehat{O^*}_{\phi,((r_1,t_1),(r_2,t_2))}(\varphi)$ (4.4), with $\phi = \pi/32$, for the spatio-temporal point configuration in Figure 4.1a (black dotted line) together with their fifth-largest and smallest envelope values (dashed lines) based on 199 random simulations of a stationary and isotropic Poisson process; the grey line is the probability value of this orientation function under the hypothesis of isotropy (which equals the uniform distribution on $[0, \pi]$, i.e. one divided by the number of angle intervals, $1/18 \simeq 0.055$).*

the same spatio-temporal short-range orientation mechanism, but now with $a = 0.5$. Visual inspection of the spatio-temporal point structure and related spatial positions of the 264 resulting points do not provide much information about the spatial structure of this point pattern (Figures 4.3a,b). Moreover, point establishment through time is quite constant and similar to the spatio-temporal point intensity $\rho = 250$ (Figures 4.3c,d). Inspection of the resulting orientation function $\widehat{O^*}_{\phi,((r_1,t_1),(r_2,t_2))}(\varphi)$ for the same spatio-temporal intervals considered in the previous scenario (see Figure 4.4) only highlights significant departures from isotropy at $\varphi = \pi/2$ for the spatio-temporal scale considered to generate this anisotropic Poisson scenario, i.e. $r_1 = 0$, $r_2 = 0.2$, $t_1 = 0$ and $t_2 = 0.1$ (see Figure 4.4a).
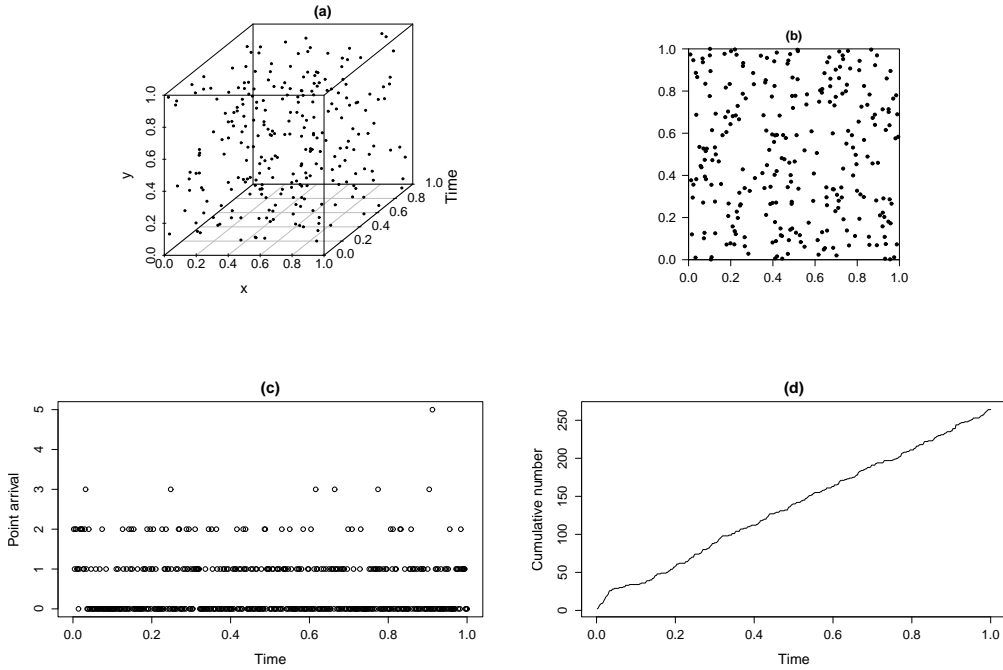
115

Figure 4.3: *As in Figure 4.1, but with $a = 0.5$.*

The rest of spatio-temporal combinations do not reveal anisotropic effects (Figures 4.4b,c,d). Under this scenario in which the anisotropic effects are not self-evident (i.e. weak anisotropic effects), our orientation function do detect anisotropic effects only for the spatio-temporal scale in which these effects are generated.

Let us now analyze the scenario with middle-range orientation effects. Figure 4.5 shows the resulting spatio-temporal point configuration for this scenario with $a = 0.5$. Here we considered $a = 0.5$ to avoid self-evident spatio-temporal anisotropic effects. Visual inspection of this spatio-temporal configuration does not provide much information about this spatial structure (with $n = 267$) (Figures 4.5a,b). Once again, time arrival and the cumulative arrival of points (Figures 4.5c,d) highlight that point establishment through time is quite constant and equal to the spatio-temporal point intensity $\rho = 250$. Figure 4.6 shows the empirical orientation function $\widehat{O^*}_{\phi,((r_1,t_1),(r_2,t_2))}(\varphi)$ for several spatio-temporal intervals (equal to those considered for the short-range orientation scenarios) and only reveals sig-
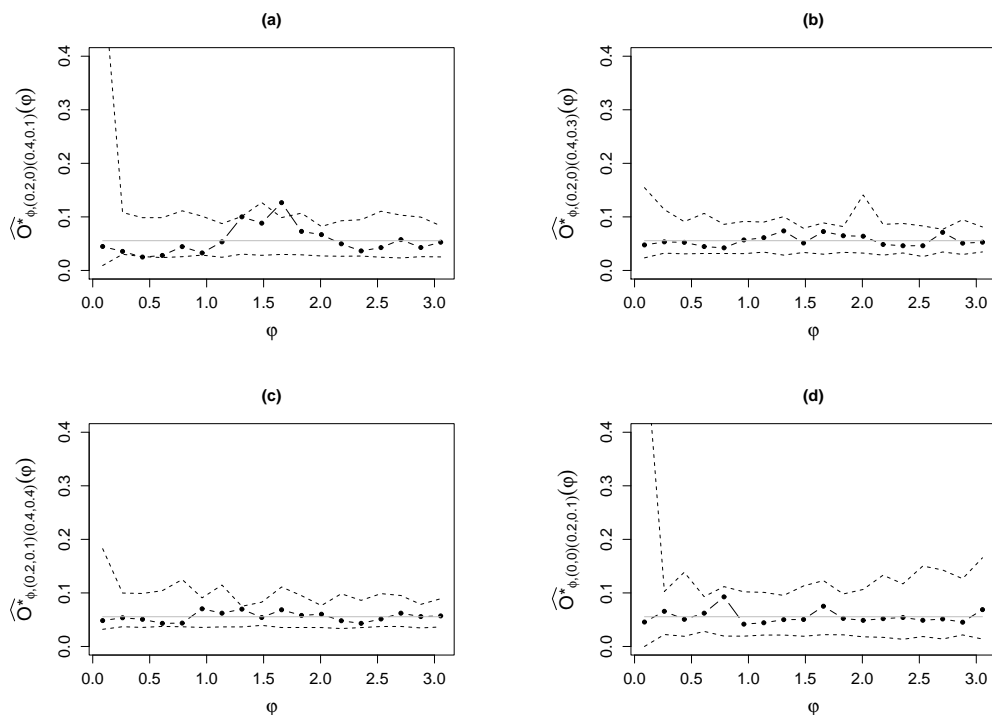
Figure 4.4: *As in Figure 4.2, but for the spatio-temporal point configuration in Figure 4.3a*

nificant anisotropic effects for the spatio-temporal interval used to generate anisotropy in the scenario under analysis, i.e. $r_1 = 0.2$, $r_2 = 0.4$, $t_1 = 0$ and $t_2 = 0.1$.

4.5 Case study: fire ignitions

We now illustrate the use of our orientation function for a real spatio-temporal point configuration. The data set is located in the northwest of the Iberian Peninsula (Spain) and provides the locations of 711 fire ignition points caused by humans located in a square area of 30×30 km for four years (1460 days), with 110 ignitions in 2007, 138 in 2008, 216 in 2009 and 247 in 2010 (see Figure 4.7). This data set was provided by the Spanish Forest Service of the Ministry of Environment and Rural and Marine Affairs. The period of study was restricted to four years due to
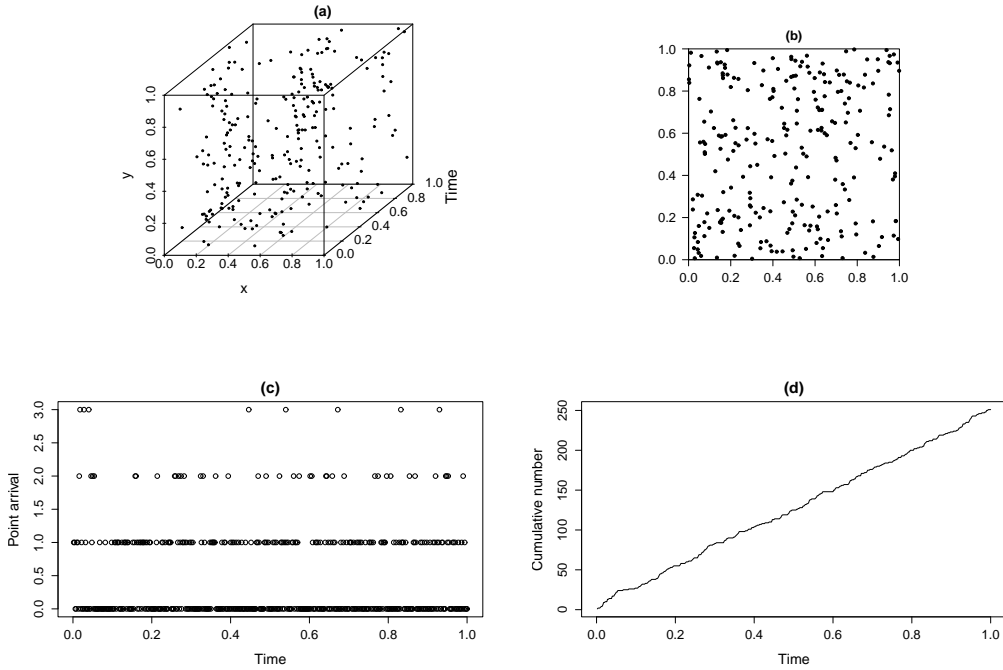
Figure 4.5: *(a) Spatio-temporal point configuration for the spatial middle-range orientation scenario ($r_1 = 0.2$, $r_2 = 0.4$, $t_1 = 0$ and $t_2 = 0.1$), and anisotropic directional effects $\beta = \pi/2$ and $c_\beta = \pi/9$ with $a = 0.5$; (b) Spatial positions of points; (c) Time arrival of points, and (d) Cumulative arrival of points.*

data availability, but this period was considered appropriate because it is the usual time framework for fire prevention planning in Spain.

Figures 4.8a,b show the spatio-temporal configuration and the spatial positions, respectively, for the 711 ignition points from 2007 to 2010, in the study area. Once again visual inspection of this spatio-temporal point configuration does not provide much information about the possible directional components of this point pattern. Time arrival and the cumulative arrival of points are shown in Figures 4.8c,d. For the time arrival of points, we also provide a parametric representation of this point sequence based on a restricted cubic spline regression (HARRELL, 2001). We used a spline regression because it is a smooth, flexible curve that makes no strict mathematical assumption on the shape of this intensity while providing a robust estimator for this data set. In particular, we used 6 knots at the 5th, 23rd,
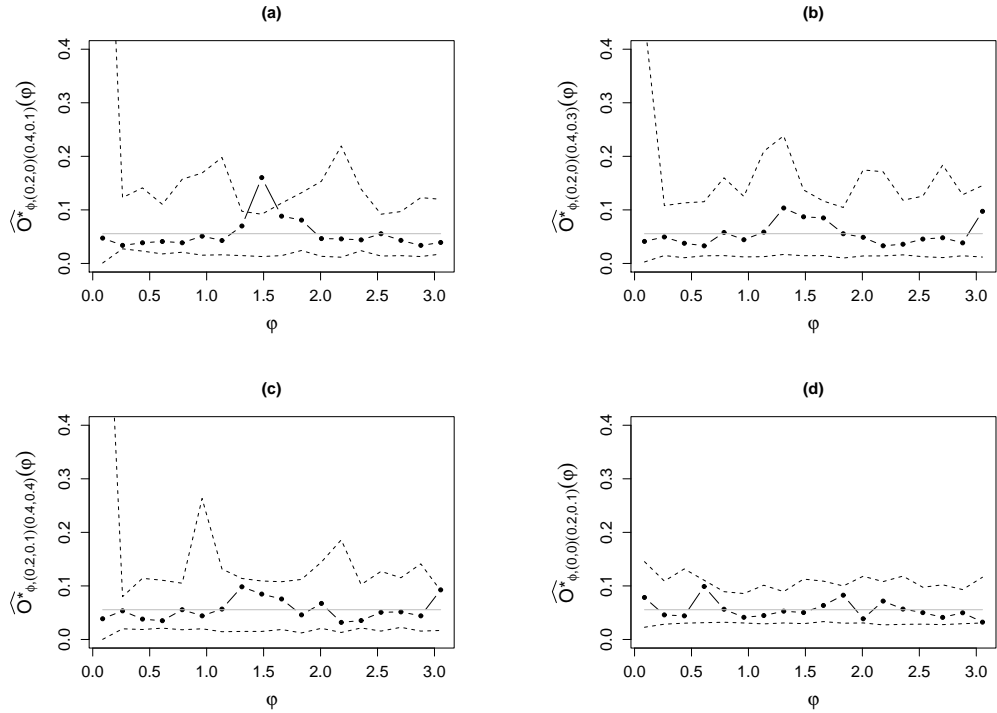
Figure 4.6: *As in Figure 4.2, but for the spatio-temporal point configuration in Figure 4.5a*

41st, 59th, 77th and 99th percentiles of these four years to provide a flexible parametric model (HARRELL, 2001). This parametric representation suggests that point arrival (establishment) is quite constant through time, as this function ranges from 0 to 1 point established per unit time for the four years of study. The cumulative point arrival also shows that point establishment is roughly constant through time. Both results confirm that for our case study the time point intensity can be assumed constant and equal to $711/(365 \times 4) \simeq 0.5$ points per day in the whole area.

We assumed two basic spatio-temporal intervals to test for spatio-temporal anisotropic effects. We considered our orientation function for short and middle spatial ranges and several time intervals. Our intention is to test the main spatio-temporal interval combinations to detect all possible directional effects. We took $r_1 = 0$ and $r_2 = 6$ km to test for short spatial ranges, and $r_1 = 6$ and $r_2 = 12$ km to analyze middle spatial

119
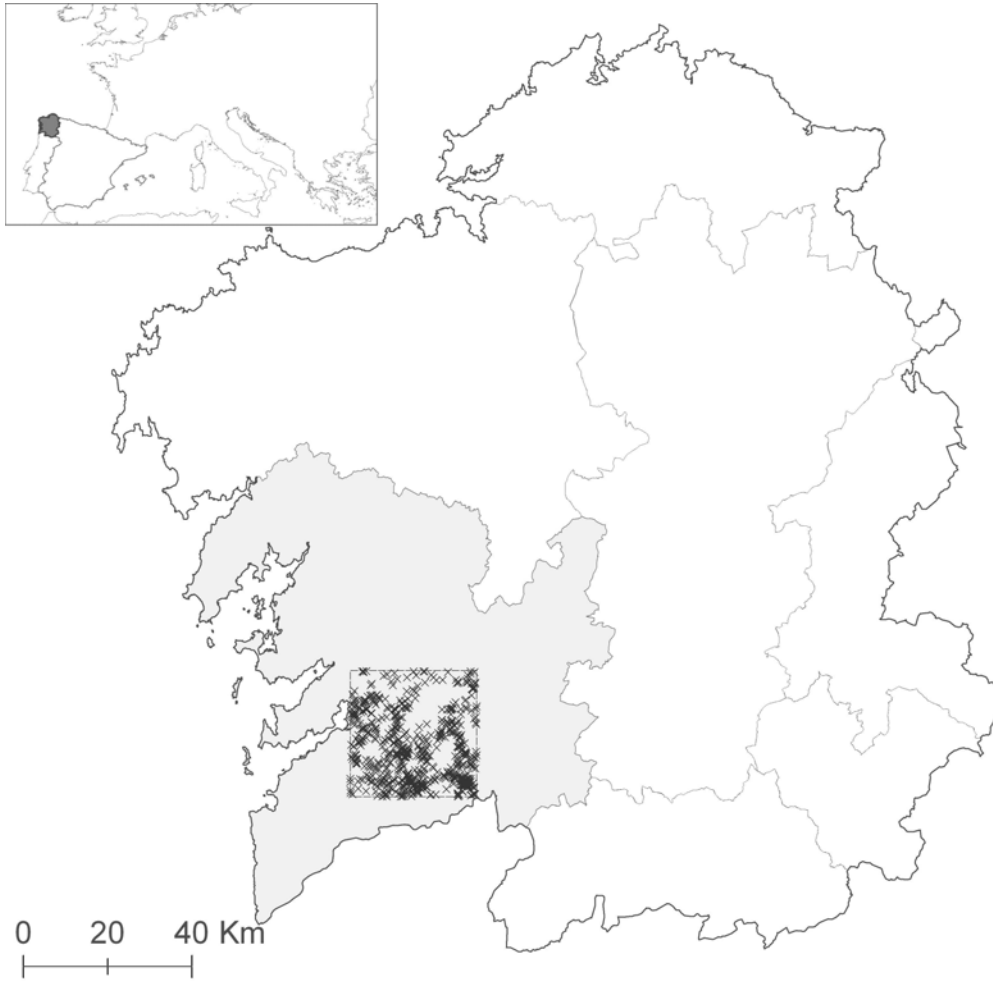
Figure 4.7: *Location of the study area, together with the* 711 *ignition points in Galicia (Spain) located inside the litoral meridional region (in grey).*

ranges. Regarding the time scales, we assumed three time intervals $t_1 = 0$ and then $t_2 = 146$ (five months) or $t_2 = 292$ (10 months) or $t_2 = 584$ ($\simeq 18$ months), which ensure an enough number of time interval combinations to scan for all time scales of anisotropy.

Figure 4.9 shows the empirical orientation function $\widehat{O^*}_{\phi,((r_1,t_1),(r_2,t_2))}(\varphi)$ (4.4) for the six spatio-temporal intervals defined above, and highlights a significant anisotropic effect at $\varphi \simeq 180^o$ for middle spatial ranges at any time interval. As this anisotropic effect loses strength when the time range

Figure 4.8: *(a) Spatio-temporal point configuration of the 711 ignition points from 2007 to 2010 in a square area located in the northw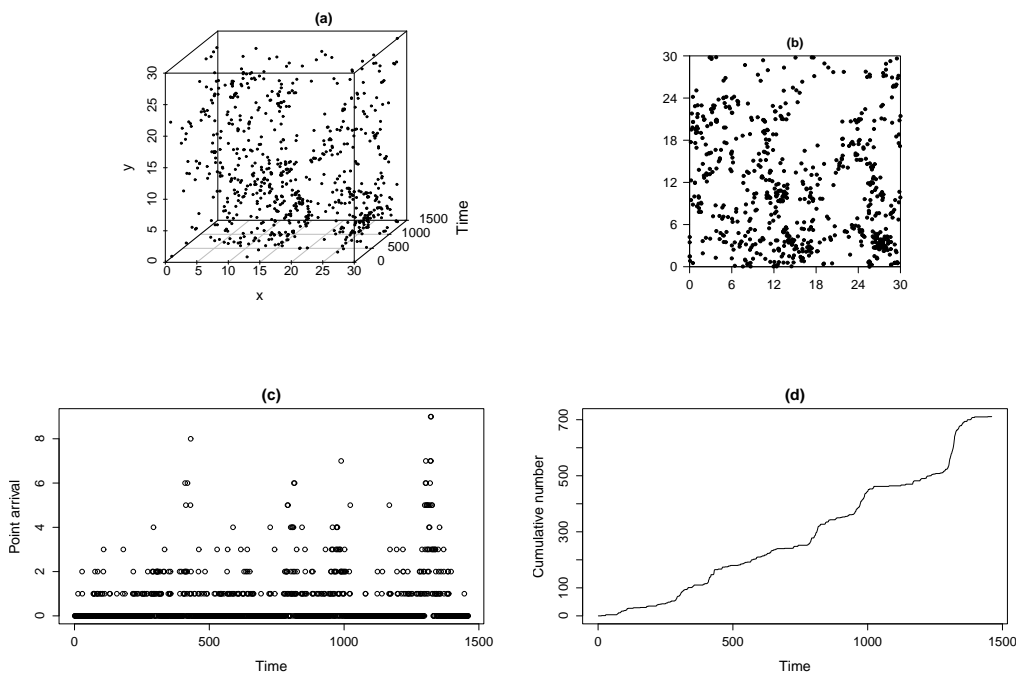est of the Iberian Peninsula; (b) Spatial positions of points; (c) Time arrival of points (grey dots) compared with fitted regression curve (black line), and (d) Cumulative arrival of points.*

interval increases from $r_2 = 146$ to $r_2 = 584$, this suggests that this directional effect happens for short time intervals (i.e. from $t_1 = 0$ to $t_2 = 146$), though it is also apparent for large time scales because $\widehat{O^*}_{\phi,((6,0),(12,584))}(\varphi)$ (say) is a cumulative function which also incorporates the directional effect detected by $\widehat{O^*}_{\phi,((6,0),(12,146))}(\varphi)$. Thus ignition points have a clear main direction of about 180° for middle spatial ranges and short time distances. This spatio-temporal anisotropic estructure (West-East direction) can be related to predominant wind directions and some geographic factors such as mountain orientation and the presence of fire barriers (routes and non-forest areas). Although it is expected that predominant winds would generate orientation effects quite independently of the spatial scale, other factor such as mountain orientation can limit such effect to some spatial

121
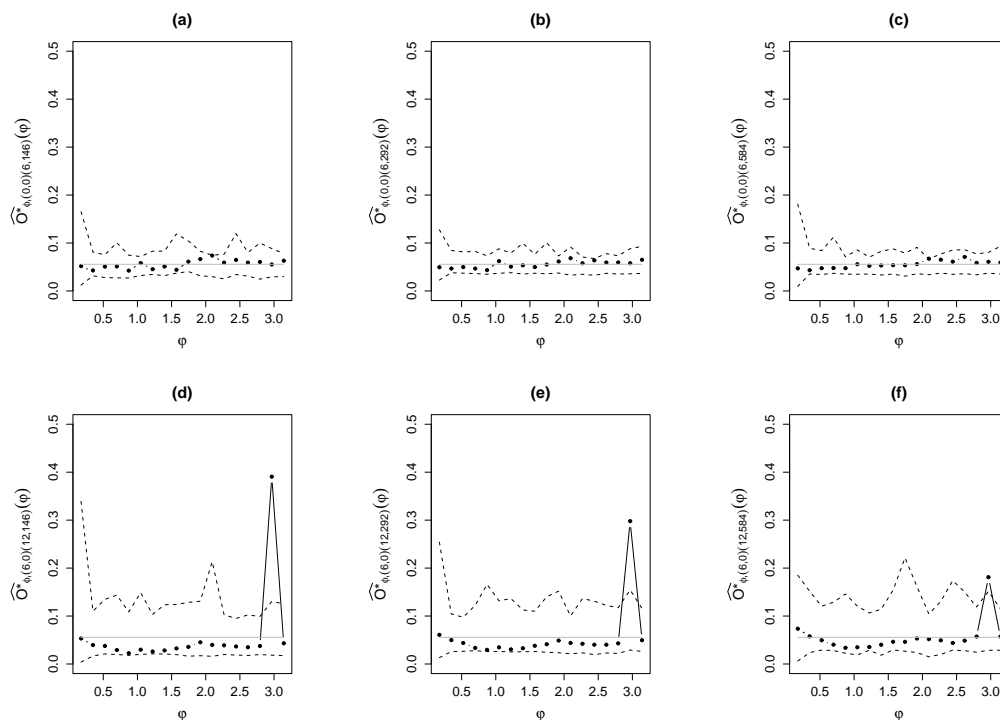
Figure 4.9: *As in Figure 4.2, but for the spatio-temporal point configuration of ignition points in Figure 4.8a*

ranges. More specific analysis should be considered incorporating some geographic covariates to fully understand the spatio-temporal orientation structure detected in this point pattern.

4.6 Discussion

We have proposed a new spatio-temporal orientation distribution function based on the spatial point pair orientation distribution function, to detect spatio-temporal anisotropic configurations. An edge-corrected estimator of this function, based on Ripley's correction, is presented and illustrated with a simulation study and with a data set involving the locations of fire ignition points caused by humans located in a square area of 30×30 km for four years. Our analysis shows that this new estimator detects spatio-temporal anisotropic structures even for weak directional effects, and it

can reveal anisotropic configurations at distinct spatio-temporal scales. In fact, our new approach is a flexible function that allows scanning for all the potentially infinite number of spatio-temporal time scales of anisotropy. Moreover, we considered a simple Monte Carlo approach to test for spatio-temporal anisotropic effects. We have followed common practice by comparing the estimator $\widehat{O^*}_{\phi,((r_1,t_1),(r_2,t_2))}(\varphi)$ with estimates obtained from simulations under a suitable null hypothesis. In this case, the null hypothesis is that the underlying point process is a stationary and isotropic Poisson process. Regarding the case study, our orientation function detects spatio-temporal anisotropic configurations not evident from the spatio-temporal point configuration. In particular, it suggests an anisotropic structure at 180° for middle spatial ranges and short time distances. This structure can probably be related to predominant wind directions and some geographic factors.

123

Chapter 5

Ongoing research

In this chapter we provide a general description of the currently ongoing research projects which have emerged motivated by the close relationship with the second-order properties of spatial and spatio-temporal point processes. We have adapted our methodology to spatio-temporal local clustering analysis and to modelling orbital debris using a new and innovative adaptation over the sphere of the classical theory.

The experience gained during the research process has allowed to gain solid foundations in the study of first- and second-order properties of spatio-temporal point processes. However it has also has generated many new questions some of which are materialised on new projects, adapting our methodology to real problems, considering new geometric contexts, and working in an interdisciplinary context. From the beginning, the objective of this thesis has focused on the generalization of the purely spatial concept into the spatio-temporal context, this seen as a hybrid of spatial and temporal components. This has been possible through extending the definition of point porcesses to include time and through the development of computational tools in R.

5.1 Local Indicators of Spatio-Temporal Association functions - LISTA functions

Francisco J. Rodríguez-Cortés[a], Mohammad Ghorbani[b] and
Jorge Mateu[a1]

[a] Department of Mathematics, Universitat Jaume I, Castellón, Spain

[b] Department of Mathematics Sciences, Aalborg University, Aalborg, Denmark

5.1.1 Introduction

Modelling real problems through spatio-temporal point processes becomes essential in many scientific and engineering fields such as environmental sciences, climate prediction and meteorology, image analysis, geology and agriculture, seismology, astronomy, epidemiology. Spatio-temporal cluster analysis is a key aspect of the practical analysis of spatio-temporal point patterns. One widely used possibility (see Chapter 3 of this thesis) is using global second-order characteristics. Moreover, in this chapter we advocate the use of local tools. The idea of considering individual contributions of a global estimator as a measure of clustering emerged in the mid 90th under the name of Local Indicators of Spatial Association (LISA), and it has been used as an exploratory data analytic tools to examine individual points in a point pattern in terms of how they relate to their neighbouring points.

These tools are based on local second-order characteristics of spatial point processes. LISA functions are built from local second-order characteristics of spatial point processes through product densities. The term LISA was coined by (Anselin (1995)), but it was later when (Cressie and Collins (2001a,b)) developed theoretical properties, namely first- and second-order moments, of these functions. Applications of LISA functions range from detecting features in images with noise (Mateu et al. (2007)) to detection of disease clusters (Moraga and Montes (2011)). Our focus

---

[1]All authors contributed equally in this work.

here is in extending the concept of LISA to the spatio-temporal context defining the LISTA functions. We define these new functions, present edge-corrected estimators, and develop their first theoretical moments. An application to detect spatio-temporal clusters in public health problems is also considered.

5.1.2 Set-up and definitions

Definitions and notations used throughout this paper are introduced by Møller and Ghorbani (2012) and Rodríguez-Cortés et al. (2014). These authors discuss the second-order analysis of structured inhomogeneous spatio-temporal point processes and provide powerful tools for the exploratory analysis of the first- and second-order characteristics. We consider a spatio-temporal point process with no multiple points as a random countable subset X of $\mathbb{R}^2 \times \mathbb{R}$, where a point $(\mathbf{u}, s) \in X$ corresponds to an event at $\mathbf{u} \in \mathbb{R}^2$ occurring at time $s \in \mathbb{R}$. In practice, we observe n events $\{(\mathbf{u}_i, s_i)\}$ of X within a bounded spatio-temporal region $W \times T \subset \mathbb{R}^2 \times \mathbb{R}$, with area $|W| > 0$, and length $|T| > 0$. Let N_{lfs} and N_{lft} be the spaces of locally finite subsets of $\mathbb{R}^2$ and $\mathbb{R}$ equipped with σ-algebras $\mathcal{N}_{\mathrm{lfs}}$ and $\mathcal{N}_{\mathrm{lft}}$ respectively, see Møller and Waagepetersen (2004). In the sequel, $N(A)$ denotes the number of the events of the process falling in a bounded region $A \subset W \times T$. For a rigurous definition of a point process based on measure theory see e.g. Daley and Vere-Jones (2008). For a given event $(\mathbf{u}, s)$, the events that are *close* to $(\mathbf{u}, s)$ in both space and time, for each spatial distance r, and time lag t, are given by the corresponding spatio-temporal cylindrical neighborhood of the event $(\mathbf{u}, s)$, which can be expressed by the cartesian product as

$$b((\mathbf{u}, s), r, t) = \{(\mathbf{v}, l) : \|\mathbf{u} - \mathbf{v}\| \le r, |s - l| \le t\}, \qquad (\mathbf{u}, s), (\mathbf{v}, l) \in \mathbb{R}^2 \times \mathbb{R},$$

where $\| \cdot \|$ denotes the Euclidean distance in $\mathbb{R}^2$ and $| \cdot |$ denotes the usual distance in $\mathbb{R}$. Note that $b((\mathbf{u}, s), r, t)$ is a cylinder with center $(\mathbf{u}, s)$, radius r and height $2t$.

5.1.3 First- and second-order properties

Assume that $\rho(\mathbf{u}, s)$ is the spatio-temporal intensity, and $\rho^{(2)}((\mathbf{u}, s), (\mathbf{v}, l))$ the second-order product density function, (see Møller and Waagepetersen (2004)). A process for which $\rho(\mathbf{u}, s) = \rho$ for all $(\mathbf{u}, s) \in X$ is called homogeneous of first-order. Further, if $\rho^{(2)}((\mathbf{u}, s), (\mathbf{v}, l)) = \rho^{(2)}(\mathbf{u} - \mathbf{v}, s - l)$, the process is called second-order or weak stationary (Ghorbani (2013)).

We assume that the point process X is orderly, roughly meaning that coincident points cannot occur. That is, any pair of points $(\mathbf{u}, s)$ and $(\mathbf{v}, l)$ of X are distinct, so $\mathbf{u} \neq \mathbf{v}$ and $s \neq l$. We can therefore ignore the case where the spatial and temporal component processes X_{space} and X_{time} have multiple points.

We assume first- and second-order spatio-temporal separability hypothesis, i.e.,

$$\rho(\mathbf{u}, s) = \bar{\rho}_1(\mathbf{u})\bar{\rho}_2(s), \quad (\mathbf{u}, s) \in \mathbb{R}^2 \times \mathbb{R}, \tag{5.1}$$

and

$$\rho^{(2)}((\mathbf{u}, s), (\mathbf{v}, l)) = \bar{\rho}_1^{(2)}(\mathbf{u}, \mathbf{v})\bar{\rho}_2^{(2)}(s, l), \quad (\mathbf{u}, s), (\mathbf{v}, l) \in \mathbb{R}^2 \times \mathbb{R} \tag{5.2}$$

where $\bar{\rho}_1, \bar{\rho}_2, \bar{\rho}_1^{(2)}, \bar{\rho}_2^{(2)}$ are non-negative functions. For more details see Møller and Ghorbani (2012) and Rodríguez-Cortés et al. (2014).

Considering the hypothesis of first-order spatio-temporal separability in (5.1), we have that

$$\rho(\mathbf{u}, s) = \frac{\rho_{\text{space}}(\mathbf{u})\rho_{\text{time}}(s)}{\int \rho(\mathbf{u}, s)\, \mathrm{d}(\mathbf{u}, s)}.$$

For a stationary point process X, $\rho, \rho_{\text{space}}$ and ρ_{time} are all constant. For non-parametric estimation of $\rho_{\text{space}}, \rho_{\text{time}}$ and $\rho(\mathbf{u}, s)$, see Ghorbani (2013).

Throughout this paper we assume that X is second-order intensity-reweighted stationary (SOIRS), i.e.

$$\rho^{(2)}((\mathbf{u}, s), (\mathbf{v}, l)) = \rho^{(2)}(\mathbf{u} - \mathbf{v}, s - l), \qquad (\mathbf{u}, s), (\mathbf{v}, l) \in \mathbb{R}^2 \times \mathbb{R}$$

(Baddeley et al.; 2000; Gabriel and Diggle; 2009). Further, if the process is

isotropic, then $\rho^{(2)}(\mathbf{u}-\mathbf{v}, s-l) = \rho_0^{(2)}(\|\mathbf{u}-\mathbf{v}\|, |s-l|)$ for some non-negative function $\rho_0^{(2)}(\cdot)$.

Just as in the spatio-temporal first-order case, considering the hypothesis of second-order spatio-temporal separability in (5.2), we have that

$$\rho^{(2)}((\mathbf{u}, s), (\mathbf{v}, l)) = \frac{\rho_{\text{space}}^{(2)}(\mathbf{u} - \mathbf{v})\rho_{\text{time}}^{(2)}(s - l)}{\int \int \rho^{(2)}(\mathbf{u} - \mathbf{v}, s - l)\,\mathrm{d}(\mathbf{u}, s)\,\mathrm{d}(\mathbf{v}, l)}. \qquad (5.3)$$

For an unbiased estimator of (5.3) and its properties of the second-order spatio-temporal product density function, see Rodríguez-Cortés et al. (2014).

5.1.4 Palm distribution

For a spatio-temporal point process $X \subset \mathbb{R}^2 \times \mathbb{R}$ with intensity measure μ (see Møller and Ghorbani (2012)), the Campbell measure is defined by the relation

$$C(B \times F) = \mathbb{E}\mathbf{1}[X \in F]N(B), \qquad B \subset \mathbb{R}^2 \times \mathbb{R}, \qquad F \in \mathcal{N}_{\text{lfs}} \times \mathcal{N}_{\text{lft}},$$

where $B = W \times T$ and $\mathbf{1}[\cdot]$ denotes the indicator function (see Møller and Waagepetersen (2004)). The reduced Campbell measure $C^!$ on $(\mathbb{R}^2 \times \mathbb{R}) \times (\mathcal{N}_{\text{lfs}} \times \mathcal{N}_{\text{lft}})$ is defined by

$$C^!(D) = \mathbb{E} \sum_{(\mathbf{u}, s) \in X} \mathbf{1}[((\mathbf{u}, s), X \setminus \{(\mathbf{u}, s)\}) \in D].$$

Analogously we can define the distribution $P_{(\mathbf{u},s)}^!$ called reduced Palm distribution and it can be interpreted as the conditional distribution of a point process given that $(\mathbf{u}, s)$ is a point of the process. For a spatio-temporal point process X and any non-negative Borel function h,

$$\mathbb{E} \sum_{(\mathbf{u},s) \in X} h((\mathbf{u}, s), X \setminus \{(\mathbf{u}, s)\})$$
$$= \int_{\mathbb{R}^2 \times \mathbb{R}} \int_{N_{\text{lfs}} \times N_{\text{lft}}} h((\mathbf{u}, s), (x, \zeta)) P_{(\mathbf{u},s)}^!(\mathrm{d}(x, \zeta))\,\mu(\mathrm{d}(\mathbf{u}, s)),$$

128

where $(x, \zeta) \in N_{\mathrm{lfs}} \times N_{\mathrm{lft}}$. In the following, $\mathbb{E}^![\cdot]$ implies expectation with respect to the reduced Palm distribution. For a stationary spatio-temporal point process X with intensity ρ and for any non-negative Borel function h

$$\mathbb{E}^! \sum_{(\mathbf{u},s) \in X} h((\mathbf{u}, s), X \setminus \{(\mathbf{u}, s)\})$$

$$= \rho \int\limits_{\mathbb{R}^2 \times \mathbb{R}} \int\limits_{N_{\mathrm{lfs}} \times N_{\mathrm{lft}}} h((\mathbf{u}, s), (\mathbf{u} + x, s + \zeta)) P_o^!(\mathrm{d}(x, \zeta)) \, \mathrm{d}(\mathbf{u}, s),$$

(5.4)

where $(\mathbf{u} + x, s + \zeta) = \{(z, \eta) + (\mathbf{u}, s) : z \in x \text{ and } \eta \in \zeta\}$ denote the translation of the point configuration (x, ζ) by $(\mathbf{u}, s)$. For more details see Daley and Vere-Jones (2003); Møller and Waagepetersen (2004).

5.1.5 Global spatio-temporal estimator

For a SOIRS and isotropic spatio-temporal point process X, Gabriel and Diggle (2009) extended the Ripley's K-function to the spatio-temporal inhomogeneous K-function. Møller and Ghorbani (2012) define the spatio-temporal inhomogeneous K-function in a more rigorously form as

$$K(r, t) = \int \mathbf{1} \left[\|\mathbf{u}\| \le r, |s| \le t \right] g_0(\mathbf{u}, s) \mathrm{d}(\mathbf{u}, s), \qquad r > 0, \quad t > 0,$$

where $g_0(\mathbf{u}, s)$ (where we abuse the notations $\mathbf{u}$ and s for $\mathbf{u} = \|\mathbf{u} - \mathbf{v}\|$ and $s = |s - l|$) is the spatio-temporal pair correlation function, where we set $a/0 = 0$ for $a \ge 0$. For a Poisson process, $g_0 = 1$ and $K(r, t) = 2\pi r^2 t$. For an unbiased estimator of the K-function (see e.g., Gabriel (2013)).

Both in the stationary and isotropic case and, the SOIRS and isotropic case, the second-order spatio-temporal product density function is proportional to the derivative of $K(r, t)$ with respect to r and t, i.e. in the planar case,

$$\rho^2(r,t) = \frac{\rho(\mathbf{u}, s)\rho(\mathbf{v}, t)}{4\pi r} \frac{\partial^2}{\partial r \partial t} K(r, t).$$

For a stationary point process X, $\rho(\mathbf{u}, s) = \rho$ and

$$\rho^{(2)}(r,t) = \frac{\rho^2}{4\pi r} \frac{\partial^2}{\partial r \partial t} K(r, t),$$

where $\rho^2 K(r,t)$ is the expected number of ordered pairs of distinct points per unit volume of the observation window with pairwise distance and time lag less than r and t, (see Rodríguez-Cortés et al. (2014)).

Under the stationarity case and ignoring edge-effects, a global naive non-parametric kernel estimator for $\rho^{(2)}(r,t)$ in (5.3) is given by

$$\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t) = \frac{1}{4\pi r |B|} \sum_{i=1}^{n} \sum_{j \neq i} \kappa_{\epsilon,\delta}(\|\mathbf{u}_i - \mathbf{u}_j\| - r, |s_i - s_j| - t), \qquad (5.5)$$

with $r > \epsilon > 0$, $t > \delta > 0$ and $B = W \times T$. We assume that the kernel function κ has the multiplicative form $\kappa_{\epsilon\delta}(\|\mathbf{u}_i - \mathbf{u}_j\| - r, |s_i - s_j| - t) = \kappa_{1\epsilon}(\|\mathbf{u}_i - \mathbf{u}_j\| - r)\,\kappa_{2\delta}(|s_i - s_j| - t)$, where $\kappa_{2\delta}$ and $\kappa_{1\epsilon}$ are respectively kernel functions with bandwidths ϵ and δ. Both the K-function and the product density function provide a global measure of the covariance structure by summing over the contributions from each event observed in the process.

5.1.6   LISTA functions

Now we consider individual contributions to the estimated function that are analogous to the local statistics described in Anselin (1995) and called local indicators of spatial association (LISA). The LISA functions were proposed by Cressie and Collins (2001a,b) in the context of spatial point processes and are similar to the notion of individual functions found in Stoyan and Stoyan (1994). We now adapt them to the spatio-temporal setting. A product density LISTA function can be constructed in the same

manner as the global estimate (5.5). It is first necessary to introduce the concept of the local version for the second-order features.

For a stationarity and isotropic spatio-temporal point process X, we can define a local version of the K-function as

$$\{\rho K(r,t)\}^i = \mathbb{E}[N(b((\mathbf{u}_i, s_i), r, t) \setminus \{(\mathbf{u}_i, s_i)\})|(\mathbf{u}_i, s_i) \in X], \quad r > 0, t > 0,$$

the expectation is conditional on observing $(\mathbf{u}_i, s_i) \in X$ and calculated with respect to the reduced Palm measure. This can be interpreted as the expected number of extra events from $(\mathbf{u}_i, s_i)$ with pairwise distance and time lag less than r and t respectively.

A spatio-temporal kernel density estimate takes the basic form of a smoothed three-dimensional histogram,

$$\frac{\rho \widehat{\partial K(r,t)}}{\partial r \partial t}^i = \sum_{j \neq i} \kappa_{\epsilon\delta}(\|\mathbf{u}_i - \mathbf{u}_j\| - r, |s_i - s_j| - t), \qquad (5.6)$$

with $r > \epsilon > 0$, $t > \delta > 0$. For a homogeneous Poisson point process $(n-1)/|B|$ provides an unbiased estimator for ρ under the reduced Palm process $X^!_{(\mathbf{u}_i, s_i)}$, since this process has the same probability distribution as the original process X. The unbiasedness is a consequence of

$$\mathbb{E}^![N(B) - 1] = \mathbb{E}\left[\int_B N^!_{(\mathbf{u}_i, s_i)}(\mathrm{d}(\mathbf{u}, s))\right] = \int_B \mathbb{E}[N(\mathrm{d}(\mathbf{u}, s))] = |B|\rho. \quad (5.7)$$

The local indicator of spatio-temporal association (LISTA) is a local function which considers individual points. The notation $\mathbb{E}^![\cdot]$ implies expectation with respect to the reduced Palm distribution (for more details see e. g. Møller and Waagepetersen (2004), Chiu et al. (2013)). A product density of the LISTA functions can be constructed in the same manner as the global estimate (5.5). We denote the localised version of the second-order product density by $\rho^{(2)i}$. Combining the kernel estimator (5.6) with

the estimate, $\widehat{\rho} = (n-1)/|B|$ a kernel estimate of $\rho^{(2)i}$ is given by

$$\widehat{\rho^{(2)i}}_{\epsilon,\delta}(r,t) = \frac{n-1}{4\pi|B|r} \sum_{i\neq j} \kappa_{1\epsilon}(\|\mathbf{u}_i - \mathbf{u}_j\| - r)\kappa_{2\delta}(|s_i - s_j| - t), \qquad (5.8)$$

with $r > \epsilon > 0$, $t > \delta > 0$. For fixed r and t it holds that

$$\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t) = \frac{1}{n-1} \sum_{i=1}^{n} \widehat{\rho^{(2)i}}_{\epsilon,\delta}(r,t),$$

which satisfies the operational definition of a LISA statistic mentioned by Anselin (1995).

5.1.7   Statistical properties

Following Cressie and Collins (2001a,b), we have that for a homogeneous Poisson process

$$\mathbb{E}^! \left[\widehat{\rho^{(2)i}}_{\epsilon,\delta}(r,t)\right] = \frac{1}{4\pi|B|r}\mathbb{E}^! \left[(N(B) - 1)\sum_{j\neq i} \kappa_{\epsilon,\delta}(\|\mathbf{u}_i - \mathbf{u}_j\| - r, |s_i - s_j| - t)\right]$$

$$= \frac{1}{4\pi|B|r}\mathbb{E}^! \left[\mathbb{E}\left\{(n-1)\sum_{j\neq i} \kappa_{\epsilon,\delta}(\|\mathbf{u}_i - \mathbf{u}_j\| - r, |s_i - s_j| - t)|N(B) = n\right\}\right]$$

$$= \frac{1}{4\pi|B|r}\mathbb{E}^! \left[(N(B) - 1)\rho \int_B \kappa_{\epsilon,\delta}(\|\mathbf{u}_i - \mathbf{x}\| - r, |s_i - \zeta| - t)\, \mathrm{d}(\mathbf{x},\zeta)\right]$$

$$= \frac{1}{4\pi|B|^2r}\mathbb{E}^! [(N(B) - 1)^2] \int_B \kappa_{\epsilon,\delta}(\|\mathbf{u}_i - \mathbf{x}\| - r, |s_i - \zeta| - t)\, \mathrm{d}(\mathbf{x},\zeta)$$

$$= \frac{(\mathbb{E}^! [N(B) - 1])^2 + \mathbb{E}^! [N(B) - 1]}{4\pi|B|^2r} \int_B \kappa_{\epsilon,\delta}(\|\mathbf{u}_i - \mathbf{x}\| - r, |s_i - \zeta| - t)\, \mathrm{d}(\mathbf{x},\zeta)$$

$$= \frac{\rho^2|B|^2 + \rho|B|}{4\pi|B|^2r} \int_{\mathbb{R}^2} \kappa_{1\epsilon}(\|\mathbf{z}\| - r)\mathbf{1}_W[\mathbf{z} + \mathbf{u}_i]\mathbf{1}[|\|\mathbf{z}\| - r| \leq \epsilon]\mathrm{d}\mathbf{z}$$

$$\times \int_{\mathbb{R}} \kappa_{2\delta}(|l| - t)\mathbf{1}_T[l + s_i]\mathbf{1}[|\|l\| - t| < \delta]\mathrm{d}l$$

132

$$= \frac{\rho^2 + \frac{\rho}{|B|}}{4\pi r} \int\limits_{r-\epsilon}^{r+\epsilon} \kappa_{1\epsilon}(u-r)|\partial b(\mathbf{u}_i, u) \cap W|_1 \mathrm{d}u \int\limits_{t-\delta}^{t+\delta} \kappa_{2\delta}(v-t)|\partial b(s_i, v) \cap T|_0 \mathrm{d}v,$$

here $|\cdot|_0$ and $|\cdot|_1$ are the zero- and one-dimensional Hausdorff measures respectively in $\mathbb{R}$ and $\mathbb{R}^2$. The zero-dimensional Hausdorff measure in $\mathbb{R}$ is the number of points of the set. The one-dimensional Hausdorff measure of a simple curve in $\mathbb{R}^2$ is equal to the length of the curve. Here we use the fact that, for a homogeneous Poisson process, conditional on observing $N(B) = n$ points and a point at $(\mathbf{u}_i, s_i)$, the events $(\mathbf{u}_1, s_1), \ldots, (\mathbf{u}_{i-1}, s_{i-1})$, $(\mathbf{u}_{i+1}, s_{i+1}), \ldots, (\mathbf{u}_n, s_n)$ are independent and identically distributed uniformly on $W \times T$. Note that with respect to the reduced Palm process $N^!_{(\mathbf{u}_i, s_i)}$, the ramdom variable $N(B) - 1$ is Poisson with mean $|B|\rho$, so that $\mathbb{E}^![(N(B) - 1)^2] = |B|^2\rho^2 + |B|\rho$.

Spatio-temporal edge-effects

A gobal edge-corrected kernel estimator of the second-order product density function is given by

$$\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t) = \frac{1}{4\pi r|B|} \sum_{i=1}^{n} \sum_{j\neq i} \frac{\kappa_{\epsilon,\delta}(\|\mathbf{u}_i - \mathbf{u}_j\| - r, |s_i - s_j| - t)}{w^2(\mathbf{u}_i, \mathbf{u}_j)w^1(s_i, s_j)},$$

with $r > \epsilon > 0$, $t > \delta > 0$, where $w^2(\mathbf{u}_i, \mathbf{u}_j)$ is the Ripley's isotropic edge-correction factor (Ripley; 1976, 1977), $w^1(s_i, s_j)$ is the temporal edge-correction factor which is equal to one if both ends of the interval of length $2|s_i - s_j|$ and center s_i lie within T, and two otherwise, Diggle et al. (1995).

Therefore, the corresponding edge-corrected second-order product density LISTA functions are

$$\widehat{\rho^{(2)i}}_{\epsilon,\delta}(r,t) = \frac{n-1}{4\pi|B|r} \sum_{i\neq j} \frac{\kappa_{1\epsilon}(\|\mathbf{u}_i - \mathbf{u}_j\| - r)\kappa_{2\delta}(|s_i - s_j| - t)}{w^2(\mathbf{u}_i, \mathbf{u}_j)w^1(s_i, s_j)}, \qquad (5.9)$$

with $r > \epsilon > 0$, $t > \delta > 0$, for $(\mathbf{u}_i, s_i) \in W \times T$ and $i = 1, \ldots, n$. They

have expected value

$$\mathbb{E}^{!}\left[\widehat{\rho^{(2)i}}_{\epsilon,\delta}(r,t)\right] = \frac{\rho^2 + \frac{\rho}{|B|}}{4\pi r}\int_{r-\epsilon}^{r+\epsilon}\frac{2\pi u}{|\partial b(\mathbf{u}_i,u)\cap W|_1}\kappa_{1\epsilon}(u-r)|\partial b(\mathbf{u}_i,u)\cap W|_1\mathrm{d}u$$

$$\times\int_{t-\delta}^{t+\delta}\frac{2}{|\partial b(s_i,v)\cap T|_0}\kappa_{2\delta}(v-t)|\partial b(s_i,v)\cap T|_0\mathrm{d}v$$

$$= \frac{\rho^2 + \frac{\rho}{|B|}}{4\pi r}\int_{r-\epsilon}^{r+\epsilon}2\pi u\kappa_{1\epsilon}(u-r)\mathrm{d}u\int_{t-\delta}^{t+\delta}2\kappa_{2\delta}(v-t)\mathrm{d}v = \rho^2 + \frac{\rho}{|B|},$$

for a homogeneous Poisson process.

The variance of the spatio-temporal edge-corrected product density LISTA can be

$$\mathbb{V}\mathrm{ar}^{!}\left[\widehat{\rho^{(2)i}}_{\epsilon,\delta}(r,t)\right] = \mathbb{E}^{!}\left[\mathbb{E}\left\{\left(\widehat{\rho^{(2)i}}_{\epsilon,\delta}(r,t)\right)^2|N(B)=n\right\}\right] - \left(\mathbb{E}^{!}\left[\widehat{\rho^{(2)i}}_{\epsilon,\delta}(r,t)\right]\right)^2$$

$$= \frac{1}{4\pi r|B|}\mathbb{E}^{!}\left[\mathbb{E}\left\{(n-1)^2\left(\sum_{i\neq j}\tau((\mathbf{u}_i,s_i),(\mathbf{u}_j,s_j),r,t)\right)^2|N(B)=n\right\}\right]$$

$$- \left[\rho^2 + \frac{\rho}{|B|}\right]^2,$$

where

$$\tau((\mathbf{u}_i,s_i),(\mathbf{u}_j,s_j),r,t) = \frac{4\pi\|\mathbf{u}_i-\mathbf{u}_j\|\kappa_{\epsilon,\delta}(\|\mathbf{u}_i-\mathbf{u}_j\|-r,|s_i-s_j|-t)}{|\partial b(\mathbf{u}_i,u)\cap W|_1|\partial b(s_i,v)\cap T|_0}.$$

5.1.8 The first computing experiences

We start by analysing the behaviour of the LISTA function on the spatio-temporal window $[0,10]^2\times[0,10]$ a spatio-temporal homogeneous Poisson process shown in Figure 5.1 with expected number of points $\mathbb{E}[N(B)] = n = 100$, and under homogeneous Poison cluster process shown in Figure 5.2 with $\rho_{\mathrm{p}} = 10$ a the intensity of the parents that follow a homogeneous Poisson process, with a number of offspring per parent being Poisson

134

with mean $m_c = 10$, and the spatial distribution of the offprints being a
zero-mean bivariate isotropic normal distribution with standard deviation
$\sigma = 2$, and a uniform temporal distribution with temporal cluster param-
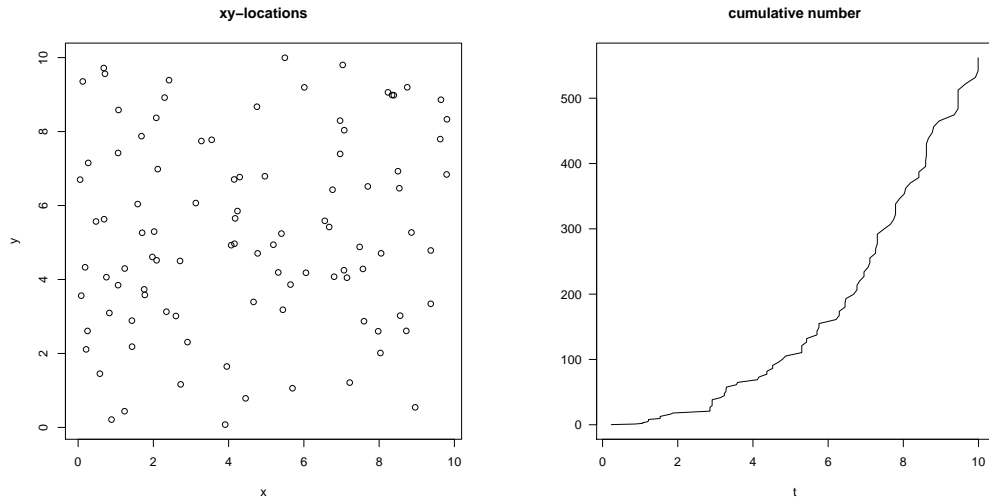eter $\delta_t = 2$.

Figure 5.1: *A spatio-temporal Poisson point pattern with expected number of the points $n = 100$.*
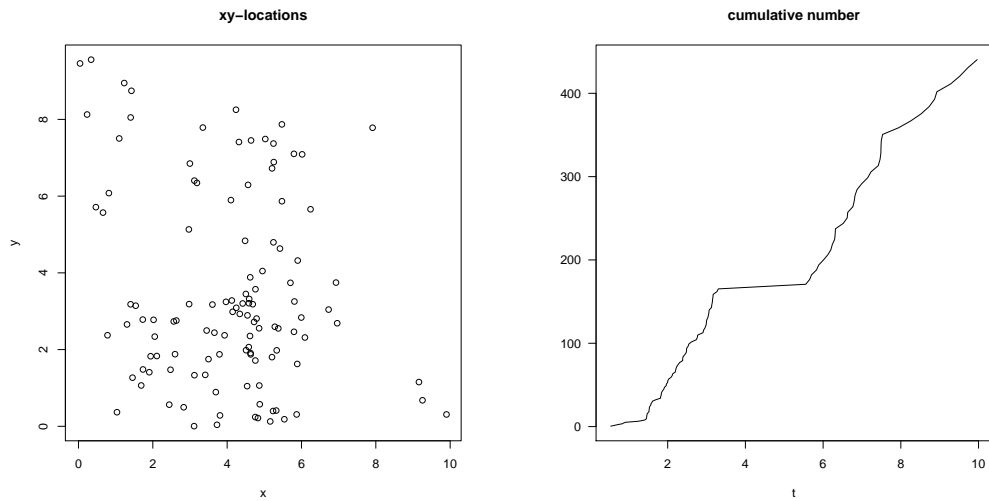

Figure 5.2: *A spatio-temporal Poisson cluster point patterns with expected number of point $n = 100$.*

We used a fine grid for each spatial and temporal distances u and v spanning the sequence starting from $u > \epsilon > 0$ to 2.50 and $v > \delta > 0$ to 2.50 with small increments of distances. In practice, we use the dpik function in kernsmooth package to obtain the bandwidth (Wand and Ripley; 2013) based on the distances between the spatial locations of the process. For the temporal case the uniform kernel is used, where again we calculate the bandwidth δ using the dpik function based on the time lag between the temporal instants of the process.
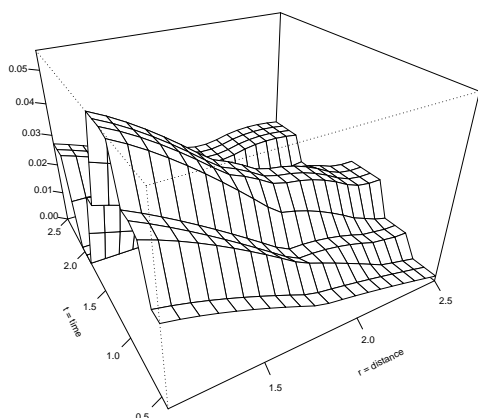
(a) LISTA for $i = 14$.
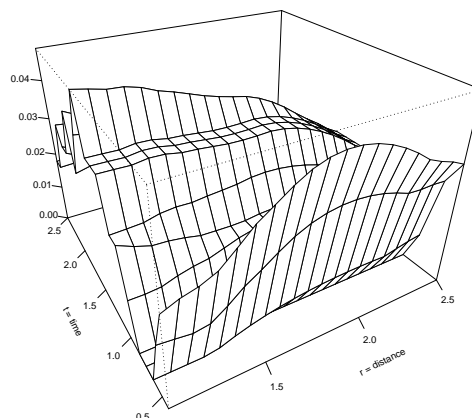
(b) LISTA for $i = 22$.
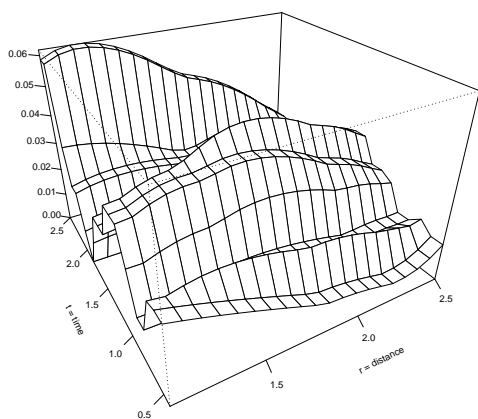
(c) LISTA for $i = 55$.

(d) LISTA for $i = 63$.

Figure 5.3: *LISTA functions kernel estimated for a homogeneous Poisson point process with $n = 100$, $\epsilon = 1.011389$ and $\delta = 0.3925461$.*

(a) LISTA for $i = 18$.

(b) LISTA for $i = 23$.
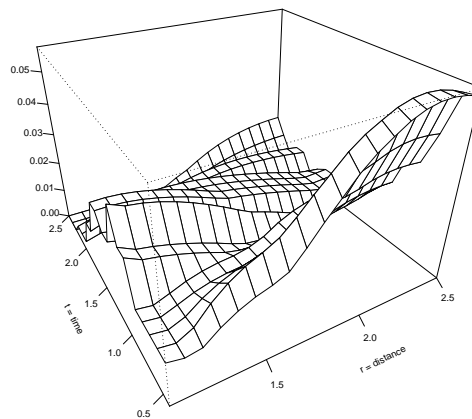
(c) LISTA for $i = 30$.

(d) LISTA for $i = 100$.
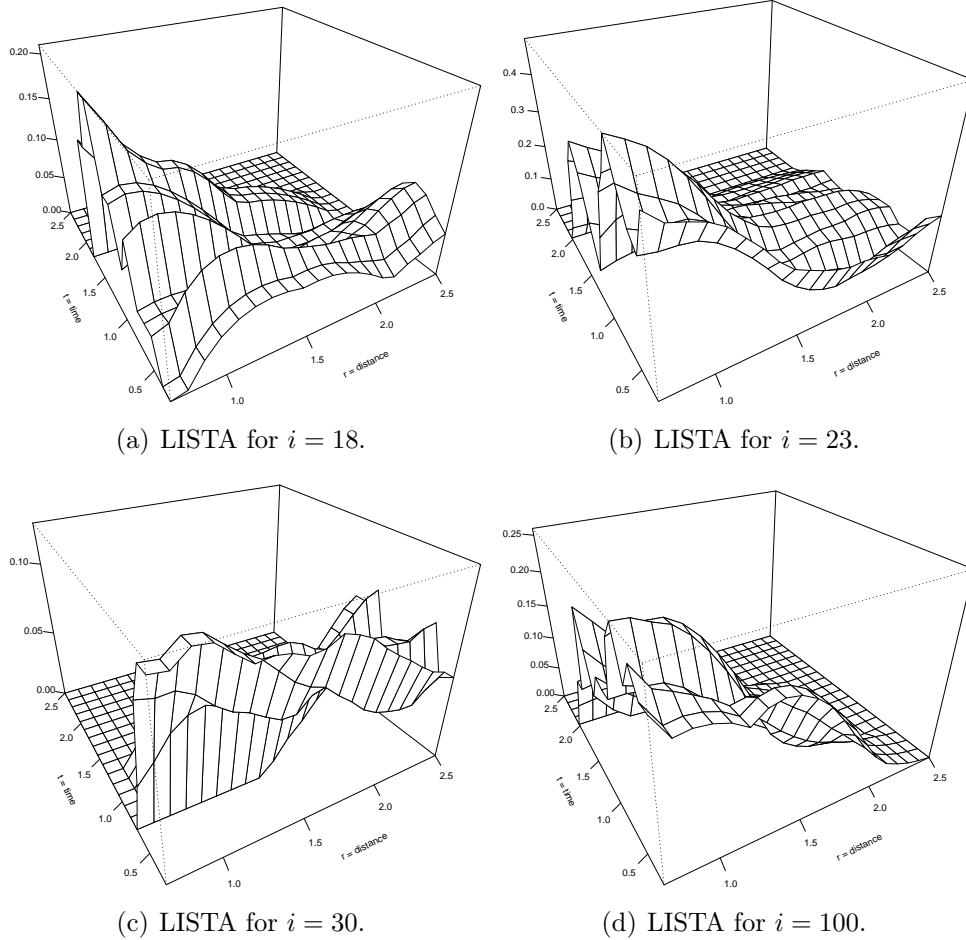
Figure 5.4: *LISTA functions kernel estimated for a homogeneous Poisson cluster point process with $n = 100$, $\epsilon = 0.6566714$ and $\delta = 0.2312543$.*

Figure 5.3 shows four randomly selected LISTA functions for the Poisson point pattern in Figure 5.1. In general the maximum value of the LISTA function surfaces is less than $\widehat{\rho^{(2)i}}_{\epsilon,\delta}(r,t) = 0.0099$ which is the estimated theoretical value under Poisson point process. Figure 5.4 shows four randomly selected LISTA functions for the Poisson cluster point pattern in the Figure 5.2. It can be seen clearly that for small spatial and temporal distances larger values of $\widehat{\rho^{(2)i}}_{\epsilon,\delta}(r,t) = 0.0099$ are obtained, which is a clear indication of the spatio-temporal cluster phenomenon.

137

5.1.9 Coming goals

- Find a closed form for the approximation of the variance of LISTA functions under the assumption of stationarity.

- Find an expression for the covariance between two LISTA functions and treatable closed forms for these expressions.

- Develop methodology for spatio-temporal clustering analysis.

- Run an intensive simulation study to test the properties of this methodology and use real data.

5.2 Modelling orbital debris with point processes on the sphere

Francisco J. Rodríguez-Cortés[a], Radu Stoica[b,c], Florent Deleflie[c]
and **Jorge Mateu[a2]**

[a] Department of Mathematics, University Jaume I, Castellón, Spain

[b] Laboratory Paul Painleve, University Lille 1, Lille, France

[c] Institute of Celestial Mechanics and Calculation of Ephemeris, Paris
Observatory, Paris, France

5.2.1 Introduction

Problems coming from the debris problem need to consider the sphere as
a metric space. In the unit sphere, we can define a probability measure
and extend the theory of point porcesses into the non-Euclidiam geome-
try context. Nowadays fields such as astronomy, engineering, agronomy,
climate prediction and meteorology are interested in this metodology.

The orbital debris often are produced by the collision of the old satel-
lites whose orbits coincide. Many theories generally accept that the Earth
is a spheric object and if we consider the effects of the Earth's gravitational
field on the orbital debris, the hypothesis that these particles are randomly
distributed around the earth is feasible.

Daley and Vere-Jones (2008) proposed a general systematic study of the
point processes and random measures that are invariant under rotations
as well as shifts in d-dimensional Euclidean space $\mathbb{R}^d$. In addition, they
comment that for processes on other types of manifolds, such as the surface
of a sphere or a cylinder the, theory is very similar, so much that some
parts are merely variants of the corresponding theory of stationary con-
tinuous processes. They consider that the basic results can be developed
with almost equal facility for point processes on a locally compact metric
group. Møller and Waagepetersen (2004) comment that in the case of the
d-dimensional unit sphere the more natural metric is given by the geodesic

---

[2] All authors contributed equally in this work.

distance. Billiot and Goulard (2001) introduced an estimation method for pairwise interaction potentials of a stationary Gibbs point process by considering the case of observations located on a sphere. This methodology is used to study independent observations of root locations on internodes around stem of maize roots.

We intend to build the foundations on a new geometric environment for the inference and estimation of the stationary point process in which the connection between the differential geometry and probability can provide to the astronomy and other sciences powerful tools to study spherical phenomena as orbital debris. The development of efficient software to perform calculations and visualization involving this methodology is one of the most challenging facts due to the high computational cost and accuracy that entails.

5.2.2   Theoretical framework

The sphere

The unit sphere S^2 with center in the origin is defined as the set of all points $\mathbf{x} \in \mathbb{R}^3$ such that

$$S^2 = \left\{ \mathbf{x} : \|\mathbf{x}\|^2 = \sum \mathbf{x}_i^2 = 1 \right\},$$

where $\| \cdot \|$ denotes the Euclidean distance in $\mathbb{R}^3$. In other words, the unit sphere S^2 is a two-dimensional closed surface embedded in three-dimensional Euclidean space.

A parametric representation of the unit sphere is given by

$$\mathbf{x}(\phi, \varphi) = (\cos(\varphi)\cos(\phi), \sin(\varphi)\cos(\phi), \sin(\phi)), \quad -\frac{\pi}{2} \leq \phi \leq \frac{\pi}{2}, \quad 0 \leq \varphi \leq 2\pi.$$

This coordinate system is especially used in astronomy for determining the latitude and longitude of the points on the globe (see Do Carmo (1976)).

Great-circle distance

Let (ϕ_i, φ_i) and (ϕ_j, φ_j) be the geographical latitude and longitude of two points on the unit sphere and

$$\Delta\phi_{ij} = |\phi_i - \phi_j|, \quad \Delta\varphi_{ij} = |\varphi_i - \varphi_j|, \quad -\frac{\pi}{2} \le \phi \le \frac{\pi}{2}, \quad 0 \le \varphi \le 2\pi,$$

their respective absolute differences, then the central angle between them $\Delta\psi$ is precisely the great-circle distance d_g among (ϕ_i, φ_i) and (ϕ_j, φ_j), that is to say d_g is the length of the shortest arc on the unit sphere between any two points. Therefore a closed expression for calculation of this distance given by the spherical law of cosines take the form

$$\Delta\psi = \arctan\left(\frac{\sqrt{\Delta\eta}}{\Delta\nu}\right), \tag{5.10}$$

where

$$\Delta\eta = (\cos(\phi_j)\sin(\Delta\varphi_{ij}))^2 + (\cos(\phi_i)\sin(\phi_j) - \sin(\phi_i)\cos(\phi_2)\cos(\Delta\varphi_{ij}))^2$$

and

$$\Delta\nu = \sin(\phi i)\sin(\phi_j) + \cos(\phi_i)\cos(\phi_j)\cos(\Delta\varphi_{ij}),$$

then $d_g((\phi_i, \varphi_i), (\phi_j, \varphi_j)) = \Delta\psi$ for more details (see Shumaker and Sinnott (1984), Vincenty (1975)).

5.2.3   Point processes on the sphere: general setting

Under similar considerations that in the planar case, we can define in point processes on the unit sphere. Let $\mathcal{B}(S^2)$ be the Borel σ-algebra in S^2 and $\mathcal{B}_0 \subseteq \mathcal{B}(S^2)$ be the system of all bounded Borel sets. We can define the space of locally finite subsets of S^2 as

$$S_{\mathrm{lf}} = \left\{x \subseteq S^2 : n(x) < +\infty, \ \forall x \in \mathcal{B}_0\right\},$$

where $n(y)$ denotes the cardinality of the set y. Elements of S_{lf} are called locally finite point configurations in the unit sphere. We equip S_{lf} with the σ-algebra

$$\mathcal{S}_{\mathrm{lf}} = \sigma\left\{\{x \in S_{\mathrm{lf}} : n(x) = m\}, m \in \mathbb{N}_0, x \in \mathcal{B}_0\right\},$$

where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. A point process on the unit sphere is also defined as a random locally finite point configuration.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an abstract probability space. A point process X_{S} in the unit sphere is a measurable mapping

$$X_{\mathrm{S}} : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (S_{\mathrm{lf}}, \mathcal{S}_{\mathrm{lf}}).$$

The distribution of the simple point processs is a measure $P_{X_{\mathrm{S}}}$ defined on $(S_{\mathrm{lf}}, \mathcal{S}_{\mathrm{lf}})$ defined by the relation

$$P_{X_{\mathrm{S}}}(F) = \mathbb{P}(X_{\mathrm{S}} \in F) = \mathbb{P}(\{\omega \in \Omega : X_{\mathrm{S}}(\omega) \in F\}), \qquad F \in \mathcal{S}_{\mathrm{lf}}.$$

A point process on the unit sphere is said finite if $n(X_{\mathrm{S}}) < +\infty$ almost surely and we will denote the random number of points falling in the set B by $N(B) = n(X_{\mathrm{S}} \cap B)$, referring to N as a counting function. A point process on the unit sphere X_{S} is stationary if its distribution is translation invariant. A point process on the unit sphere X_{S} is isotropic if its distribution is invariant under rotations. These two concepts are given in the same sense as the planer case, (see Møller and Waagepetersen (2004)).

It is natural to think that the important properties and basic results of point processes in the plane can be immediately extended into this new context. However, our intention in this paper is only to show the how to adapt classical theory of stationary point processes to the case where is the unit sphere. The intensity measure μ_{S} on S^2 is given by

$$\mu_{\mathrm{S}}(B) = \int_B \rho_{\mathrm{S}}(\mathbf{x})\,\mathrm{d}\mathbf{x}, \qquad B \in \mathcal{B},$$

where ρ_S is a non-negative function. If ρ_S is constant, then X_S is said to be homogeneous or first-order stationary with intensity ρ_S; otherwise X_S is said to be inhomogeneous. The second-order factorial moment measure $\alpha_S^{(2)}$ on S^2 is given by

$$\alpha_S^{(2)}(C) = \iint \mathbf{1}\left[(\mathbf{x}, \mathbf{y}) \in C\right] \rho_S^{(2)}(\mathbf{x}, \mathbf{y}) \, \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}, \qquad C \in \mathcal{B} \times \mathcal{B},$$

where $\rho_S^{(2)}$ is a non-negative function and called second-order product density. Therefore, if ρ_S and $\rho_S^{(2)}$ coexist then the pair correlation function is defined by

$$g_S(\mathbf{x}, \mathbf{y}) = \frac{\rho_S^{(2)}(\mathbf{x}, \mathbf{y})}{\rho_S(\mathbf{x})\rho_S(\mathbf{y})}$$

where we consider $a/0 = 0$ for $a \geq 0$.

Suppose that X_S has intensity function ρ_S and that the measure

$$\mathcal{K}_S(B) = \frac{1}{|A|}\mathbb{E}\left[\sum_{\mathbf{x}, \mathbf{y} \in X_S}^{\neq} \frac{\mathbf{1}\left[\mathbf{y} \in A, \mathbf{x} - \mathbf{y} \in B\right]}{\rho_S(\mathbf{x})\,\rho_S(\mathbf{y})}\right], \qquad B \subseteq S^2, \qquad (5.11)$$

does not depend on the choice of $A \in \mathcal{B}$ with Lebesgue measure $|A|$. Accordingly X_S is said to be second-order intensity reweighted stationary (SOIRS) and $\mathcal{K}_S$ is called the second-order reduced moment measure. Hence every stationary point process on the unit sphere is SOIRS.

If the pair correlation function exists and is invariant under translations, then the second-order intensity reweighted stationarity $\mathcal{K}_S$ is given by

$$\mathcal{K}_S(B) = \int_B g_S(\mathbf{x})\mathrm{d}\mathbf{x}, \qquad B \subseteq S^2. \qquad (5.12)$$

5.2.4 Simulating an homogeneous Poisson point process on the unit sphere

As it is well known, Poisson processes play a key role in the development of the theory of point processes in any geometrical context. In the case of the unit sphere, the pseudo-algorithm for its generation must be modified

for avoid accumulation at the poles. This algorithm can be depicted as follow:

1. Generation of the number n of geographical points on the unit sphere S^2, where n is a sample of a Poisson distributed random variable with parameter $4\pi\rho_S$,

2. Generation of a sample of a binomial point process on the unit sphere S^2 with n geographical points:

   - Longitude: Given n, generate independent random sample from the uniform distribution on $[0, 2\pi]$.

   - Latitude: Given n, generate on independent random sample from the uniform distribution on $[-1, 1]$ and consider the transformation $\arccos(\cdot) - \pi/2$.

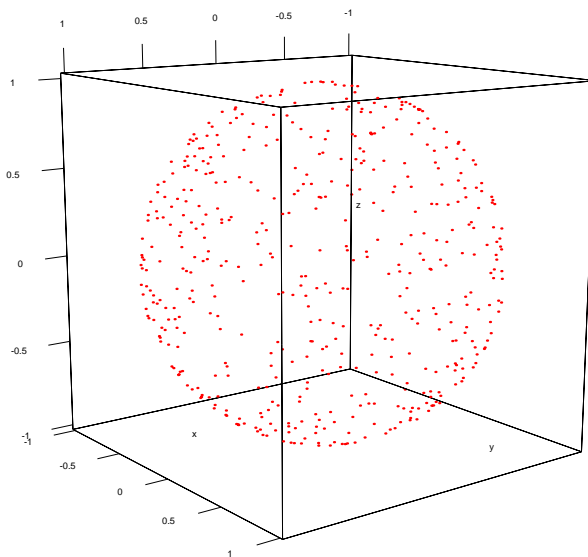for more detalles see Section 1.3.1

Figure 5.5: *Homogeneous Poisson point pattern on the unit sphere for $n = 500$.*

Figure 5.5 shows a realisation of a homogeneous Poisson point pattern on the unit sphere with $\widehat{\rho}_S = 39.78874$.

144

5.2.5 Second-order summary statistics

The K_S-function for a SOIRS point processes on the unit sphere are defined by

$$K_S(c) = \mathcal{K}_S(b(0,c)), \qquad 0 \le c \le \pi,$$

where $b(0,c)$ is a ball with form of the spherical cap with center in origin and the spherical length radius is the great-circle distance on the unit sphere define in (5.10). In the stationary case, $\rho_S K_S(c)$ is the expected number of further points within distance c from the origin given that X_S has a point at the origin, (Baddeley et al. (2000), Møller and Waagepetersen (2004)).

If g_S is isotropic, i.e. $g_S(\mathbf{x}, \mathbf{y}) = g(|\mathbf{x} - \mathbf{y}|_g)$ where $d_g(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|_g$, we can write the the K_S-function as

$$K_S(c) = \int \mathbf{1}[|\mathbf{u}|_g \le c] g_S(\mathbf{u}) \mathrm{d}\mathbf{u}, \qquad 0 \le c \le \pi,$$

were $\mathbf{1}[\cdot]$ denotes the indicator function and we abuse notation for denote by $g_S(\cdot)$ also the function which describes how the pair correlation only depends on interpoint distances.

In the isotropic case the pair correlation function has relationship through of the derivated with the K_S-function as

$$g_S(c) = \frac{1}{2\pi \sin(c)} \frac{\mathrm{d}}{\mathrm{d}c} K_S(c), \qquad 0 \le c \le \pi. \tag{5.13}$$

For a Poisson point process on the unit sphere holds that,

$$K_S(c) = 2\pi(1 - \cos(c)) \qquad \text{and} \qquad g_S(c) = 1. \tag{5.14}$$

where c is the great-circle distance.

Estimates for the K_S-function and the second-order product density function on the unit sphere are given by

$$\widehat{K_S}(c) = \sum_{\mathbf{x}, \mathbf{y} \in X_S}^{\ne} \frac{\mathbf{1}[|\mathbf{x} - \mathbf{y}|_g \le c]}{4\pi \hat{\rho}_S(\mathbf{x}) \hat{\rho}_S(\mathbf{y})}, \qquad \mathbf{x}, \mathbf{y} \in S^2, \, 0 \le c \le \pi, \tag{5.15}$$

and

$$\widehat{\rho_{\mathrm{S}}^{(2)}}_\epsilon(c) = \sum_{\mathbf{x},\mathbf{y}\in X_{\mathrm{S}}}^{\neq} \frac{\kappa_\epsilon(|\mathbf{x}-\mathbf{y}|_g - c)}{8\pi^2 \sin(c)}, \quad \mathbf{x},\mathbf{y}\in S^2,\ 0\le c\le \pi. \qquad (5.16)$$

Given that the unit sphere is a closed surface in the construction of the estimators for K_{S} and g_{S} is obvious that it is not necessary to consider edge-effect and therefore the proof of its unbiasedness of the estimators just repeat steps as the spatial case.

5.2.6 The first computing experiences

We generate of a homogeneous Poisson point pattern with expected number of the point per unit of area $\mathbb{E}[N(S^2)] = 500$, by the algorithm we have outlined in Section 5.2.4 on the unit sphere. We used a fine spherical grid of the great-circle distances c spanning the sequence starting from $c > 0$ to $\pi/4$ with small increments of the reat-circle distances on the unit sphere.
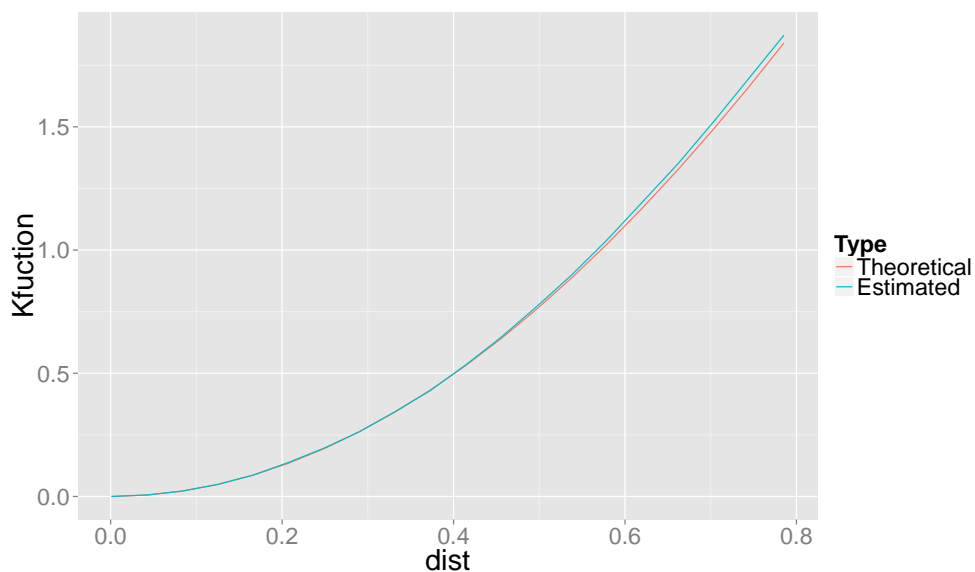

Figure 5.6: *Theoretical and estimated K-functions of a realisation of a homogeneous Poisson point pattern with $n = 500$ points in the unit sphere.*

Figure 5.6 shows the estimator of the K_{S}-function in (5.15) under a

homogeneous Poisson point pattern on the unit sphere and we note that the theoretical expression (5.14) are certainly close to the empirical one.
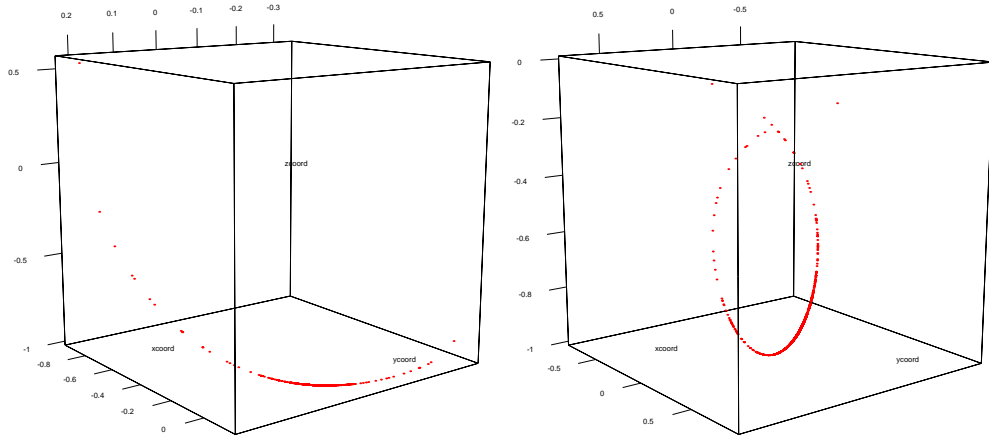

Figure 5.7: *Simulation of the behavior of particles detached by satellites moments after a collision.*
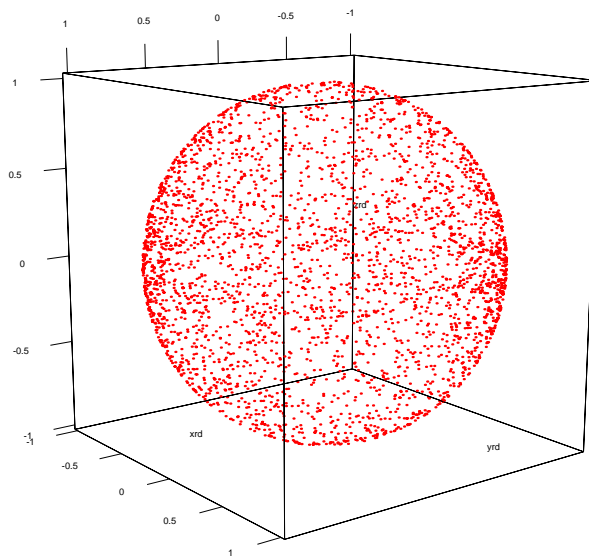

Figure 5.8: *Geographical locations of 3255 orbital debris particles orbiting the Earth.*

Figure 5.7 illustrates the spherical distribution of particles after collision. They are clustered with a strong aggregation. Figure 5.8 shows simulated data from the Observatory of Lille - France wiht Reference ellipsoid: WGS-84 under conditions of the day January 15 of 2013. This

consists of geographical locations of 3255 orbital debris particles orbiting the Earth. Astronomers expected that this cloud of orbital debris with the passage of time ends up on random distributions on unit sphere.

Figure 5.9: *95%-envelope obtained from 39 simulations of a homogeneous Poisson point pattern on the unit sphere for n = 3255.*
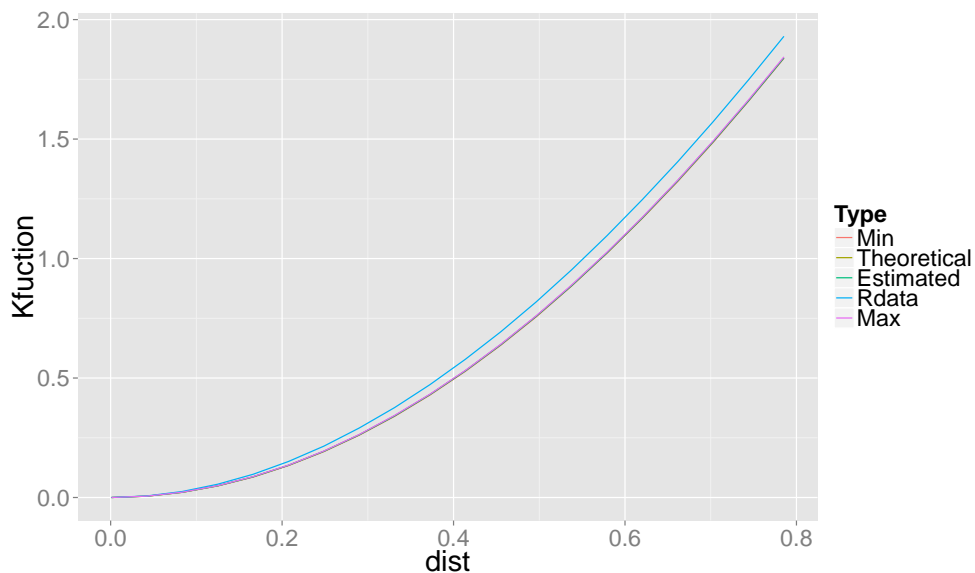
Figure 5.9 shows the 95%-envelope obtained from 39 simulations of a homogeneous Poisson point pattern on the unit sphere with $n = 3255$. Here we can see how the blue curve which represents the simulated data from particles around of the Earth, is larger than the upper 95%-envelope. Thus we can consider the data shown in Figure 5.8 are distributed around the sphere as a cluster pattern.

5.2.7 Coming goals

- Further developing the methodology of point processes in the unit sphere.

- Extension of computational tools for the statistical analysis in the context of geometry on the sphere.

- Provide answers to other scientists with arguments based on the stochastic geometry for give tools estimation and inference.

Bibliography

Adler, R. (1981). *The Geometry of Random Fields*, Wiley, New York.

Akaike, H. (1954). An approximation to the density function, *Annals of the Institute of Statistical Mathematics* **6**: 127–132.

Anselin, L. (1995). Local indicators of spatial association–LISA, *Geographical Analysis* **27**(2).

Baddeley, A., Gregori, P., Mateu, J., Stoica, R. and Stoyan, D. (eds) (2006). *Case Studies in Spatial Point Process Modeling*, Springer Lecture Notes in Statistics 185, Springer-Verlag, New York.

Baddeley, A., Møller, J. and Waagepetersen, R. (2000). Non- and semi-parametric estimation of interaction in inhomogeneous point patterns, *Statistica Neerlandica* **54**: 329–350.

Baddeley, A. and Turner, R. (1995). spatstat: An r package for analyzing spatial point patterns, *Journal of Statistical Software* **12**(06): 1–42.

Bedford, T. and van den Berg, J. (1997). A remark on van Lieshout and Baddeley's J-function for point processes, *Advances in Applied Probability* **29**: 19–25.

Berlinet, A. and Thomas-Agnan, C. (2004). *Reproducing Kernel Hilbert Spaces in Probability and Statistics*, Springer US.

Berman, M. and Diggle, P. (1989). Testing for spatial association between a point pattern and another stochastic process, *Journal of Royal Statistical Society Series B* **51**: 81–92.

Besag, J. E. (1977). Discussion of the paper by Ripley (1977), *Journal of the Royal Statistical Society Series B* **39**: 193–195.

Billiot, J.-M. and Goulard, M. (2001). An estimation method of the pair potential function for gibbs point processes on spheres, *Scandinavian journal of statistics* **28**(1): 185–203.

Brillinger, D. R. (1975). *Stochastic Processes and Related Topics*, Academic Press, New York, chapter Statistical inference for stationary point porcess, pp. 55–99.

Brix, A. and Møller, J. (2001). Space-time multitype log Gaussian Cox processes with a view to modelling weed data, *Scandinavian Journal of Statistics* **28**: 471–488.

Chiu, S. N., Stoyan, D., Kendall, W. S. and Mecke, J. (2013). *Stochastic Geometry and Its Applications*, third edn, John Wiley & Sons.

Choi, E. and Hall, P. (2001). Nonparametric approach to analysis of space-time data on earthquake occurrences, *Journal of Computational and Graphical Statistics* **8**: 733–748.

Cliff, A. D. and Ord, J. K. (1981). *Spatial Processes: Models and Applications*, Poon Limited, London.

Cline, D. B. H. (1988). Admissibile kernel estimators of a multivariate density, *Ann. Statist.* **16**: 1421–1427.

Comas, C. and Mateu, J. (2011). Statistical inference for gibbs point processes based on field observations, *Stoch Environ Res Risk Assess* **25**: 287–300.

Comas, C., Palahi, M., Pukkala, T. and Mateu, J. (2009). Characterising forest spatial structure through inhomogeneous second order characteristics, *Stoch Environ Res Risk Assess* **23**: 387–397.

Cox, D. R. (1955). Some statistical models related with series of events, *Journal of the Royal Statistical Society Series B* **17**: 129–164.

Cox, D. R. and Isham, V. (1980). *Point Processes*, Chapman & Hall, London.

Cressie, N. A. C. (1993). *Statistics for Spatial Data*, second edn, Wiley, New York.

Cressie, N. and Collins, L. B. (2001a). Analysis of spatial point patterns using bundles of product density lisa functions, *Journal of Agricultural, Biological, and Environmental Statistics* **6**: 118–135.

Cressie, N. and Collins, L. B. (2001b). Patterns in spatial point locations: Local indicators of spatial association in a minefield with clutter, *Naval Research Logistics* **48**: 333–347.

Cressie, N. and Wikle, C. K. (2011). *Statistics for spatio-temporal data*, Wiley. com.

Cronie, O. and Lieshout, M. N. M. v. (2014). A J-function for inhomoge neous spatio-temporal point processes, *Submitted* .

Daley, D. J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods*, second edn, Springer-Verlag, New York.

Daley, D. J. and Vere-Jones, D. (2008). *An Introduction to the Theory of Point Processes. Volume II: General Theory and Structure*, second edn, Springer-Verlag, New York.

De Forest, E. L. (1873). *Annual report of the Board of Regents of the Smithsonian Institution*, Vol. 1871, Washington :Smithsonian Institu-

tion, chapter On some methods of interpolation applicable to the graduation of irregular series, such as tables of mortality, &c., &c, pp. 275 – 340.

Diggle, P., Chetwynd, A., Häggkvist, R. and Morris, S. (1995). Second-order analysis of space-time clustering, *Statistical Methods in Medical Research* **4**: 124–136.

Diggle, P. J. (1983). *Statistical Analysis of Spatial Point Patterns*, Academic Press, London.

Diggle, P. J. (1985). A kernel method for smoothing point process data, *Applied Statistics* **34**: 138–147.

Diggle, P. J. (2003). *Statistical Analysis of Spatial Point Patterns*, second edn, Arnold, London.

Diggle, P. J. (2006). Spatio-temporal point processes, partial likelihood, foot and mouth disease, *Statistical Methods in Medical Research* **15**: 325–336.

Diggle, P. J. (2013). *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*, CRC Press.

Diggle, P. J. and Gabriel, E. (2010). Spatio-temporal point processes, *in* A. E. Gelfand, P. J. Diggle, P. Guttorp and M. Fuentes (eds), *Handbook of Spatial Statistics*, CRC Press, Boca Raton, pp. 451–464.

Diggle, P. J., Mateu, L. and Clough, H. E. (2000). A comparison between parametric and non-parametric approaches to the analysis of replicated spatial point patterns, *Advances of Applied Probability* **32**: 331–343.

Diggle, P. J. and Richardson, S. (1993). Epidemiological studies of industrial pollutants: An introduction, *International Statistical Review / Revue Internationale de Statistique* **61**: 203–206.

Do Carmo, M. P. (1976). *Differential geometry of curves and surfaces*, Vol. 2, Prentice-Hall Englewood Cliffs.

Doguwa, S. I. (1990). On edge-corrected kernel-based pair-correlation function estimators for point processes, *Biometrical Journal* **32**: 95–106.

Donoho, D. L. (1993). Nonlinear wavelet methods for recovery of signals, densities, and spectra from indirect and noisy data, *Proceedings of symposia in Applied Mathematics*, Vol. 47, Providence: American Mathematical Society, pp. 173–205.

Efromovich, S. (1999). *Nonparametric Curve Estimation Methods, Theory, and Applications*, Springer, New York.

Epanechnikov, V. A. (1969). Non-parametric estimation of a multivariate probability density read more, *Theory of Probability & Its Applications* **14**: 153–158.

Falk, M. (1983). Relative efficiency and deficiency of kernel type estimators of smooth distribution functions, *Statistica Neerlandica* **37**: 73–83.

Fan, J. and Marron, J. S. (1994). Fast implementations of nonparametric curve estimators, *Journal of Computational and Graphical Statistics* **3**: 35–56.

Fiksel, T. (1988a). Edge-corrected density estimators for point processes, *Statistics* **19**: 67–75.

Fiksel, T. (1988b). Estimation of interaction potentials of gibbsian point processes, *Statistics* **19**: 77–86.

Gabriel, E. (2013). Estimating second-order characteristics of inhomogeneous spatio-temporal point processes, *Methodology and Computing in Applied Probability* pp. 1–21.

Gabriel, E. and Diggle, P. J. (2009). Second-order analysis of inhomogeneous spatio-temporal point process data, *Statistica Neerlandica* **63**: 43–51.

Gabriel, E., Rowlingson, B. and Diggle, P. J. (2012). Stpp: Space-time point pattern simulation, visualisation and analysis. R package version 0.2.

Gabriel, E., Wilson, D. J., Leatherbarrow, A. J., Cheesbrough, J., Gee, S., Bolton, E., Fox, A., Fearnhead, P., Hart, C. A. and Diggle, P. J. (2010). Spatio-temporal epidemiology of campylobacter jejuni enteritis, in an area of northwest england, 2000–2002, *Epidemiology and infection* **138**(10): 1384–1390.

Gao, W. and Li, B. (1993). Wavelet analysis of coherent structures at the atmosphere-forest interface, *Journal of Applied Meteorology* **32**(11): 1717–1725.

Gasser, T. and Müller, H.-G. (1984). Estimating regression functions and their derivatives by the kernel method, *Scandinavian Journal of Statistics* **11**: 171–185.

Gasser, T., Müller, H.-G. and Mammitzsch, V. (1985). Kernels for non-parametric curve estimation, *Journal of the Royal Statistical Society. Series B (Methodological)* **47**: 238–252.

Gasser, T. and Rosenblatt, M. (1979). *Smoothing Techniques for Curve Estimation*, Springer Berlin Heidelberg, chapter Kernel estimation of regression functions, pp. 23–68.

Gelfand, A., Diggle, P., Guttorp, P. and Fuentes, M. (2010). *Handbook of Spatial Statistics*, Chapman & Hall/CRC Press.

Ghorbani, M. (2013). Testing the weak stationarity of a spatio-temporal point process, *Stochastic Environonmental Research and Risk Assessesment* **27**: 517–524.

Goodall, D. (1952). Some considerations in the use of point quadrats for the analysis of vegetation, *Australian Journal of Biological Sciences* **5**: 1 – 41.

Gram, J. (1883). Ueber die entwickelung reeller functionen in reihen mittelst der methode der kleinsten quadrate., *Journal für die reine und angewandte Mathematik* **94**: 41–73.

Grenfell, B., Bjørnstad, O. and Kappey, J. (2001). Travelling waves and spatial hierarchies in measles epidemics, *Nature* **414**(6865): 716–723.

Guan, Y. (2007). A least-squares cross-validation bandwidth selection approach in pair correlation function estimations, *Statistics & Probability Letters* **77**: 1722 – 1729.

Guan, Y., Sherman, M. and Calvin, J. A. (2004). A nonparametric test for spatial isotropy using subsampling, *Journal of the American Statistical Association* **99**(467): 810–821.

Guan, Y., Sherman, M. and Calvin, J. A. (2006). Assessing isotropy for spatial point processes, *Biometrics* **62**: 119–125.

Harper, K. and Macdonald, S. (2001). Structure and composition of riparian boreal forest: new methods for analyzing edge influence, *Ecology* **82**(3): 649–659.

Hart, J. D. (1997). *Nonparametric Smoothing and Lack-of-Fit Tests*, Springer, New York.

Hodder, I. and Orton, C. (1976). *Spatial analysis in archaeology*, Cambridge University Press.

Hoem, J. M. (1983). The reticent trio: Some little-known early discoveries in life insurance mathematics by l. h. f. oppermann, t. n. thiele and j. p. gram, *International Statistical Review* **51**: 213–221.

Illian, J. B., Møller, J. and Waagepetersen, R. P. (2009). Hierarchical spatial point process analysis for a plant community with high biodiversity, *Environmental and Ecological Statistics* **16**: 389–405.

156

Illian, J., Penttinen, A., Stoyan, H. and Stoyan, D. (2008). *Statistical Analysis and Modelling of Spatial Point Patterns*, John Wiley and Sons, Chichester.

Juan, P., Mateu, J. and Saez, M. (2012). Pinpointing spatio-temporal interactions in wildfire patterns, *Stoch Environ Res Risk Assess* **26**: 1131–1150.

Kovalev, V. A. and Bondar, Y. S. (1997). A method for anisotropy analysis of 3d images, *Computer Analysis of Images and Patterns*, Springer, pp. 495–502.

Krickeberg, K. (1982). *Ecole d'Eté de Probabilités de Saint-Flour X - 1980*, Springer, Berlin Heidelberg, chapter Processus ponctuels en statistique, pp. 205–313.

Lieshout, M. N. M. v. and Baddeley, A. J. (1996). A nonparametric measure of spatial interaction in point patterns, *Statistica Neerlandica* **50**: 344–361.

Lotwick, H. W. and Silverman, B. W. (1981). Convergence of spatial birth-and-death processes, *Math. Proc. Cambridge Philos. Soc.* **90**: 155–165.

Lotwick, H. W. and Silverman, B. W. (1982). Methods for analysing spatial point processes of several types of points, *Journal of the Royal Statistical Society Series B* **44**: 406–413.

Luke, Y. L. (1969). *The special functions and their approximations*, Academic Press, New York.

Macaulay, F. R. (1931). *The Smoothing of Time Series*, The National Bureau of Economic Research, chapter The Smoothing of Economic Time Series, Curve Fitting and Graduation, pp. 31 – 42.

Mammitzsch, V. (1984). On the asymptotically optimal solution within a certain class of kernel type estimators, *Statist. Decisions* **2**: 247–255.

Marron, J. and Nolan, D. (1988). Canonical kernels for density estimation, *Statistics & Probability Letters* **7**: 195 – 199.

Matérn, B. (1986). *Spatial Variation*, Lecture Notes in Statistics 36, Springer-Verlag, Berlin.

Mateu, J. (2000). Second-order characteristics of spatial marked processes with applications, *Nonlinear Analysis: Real World Applications* **1**: 145–162.

Mateu, J., Lorenzo, G. and Porcu, E. (2007). Detecting features in spatial point processes with clutter via local indicators of spatial association, *Journal of Computational and Graphical Statistics* **16**(4): 968–990.

Messer, K. and Goldstein, L. (1993). A new class of kernels for nonparametric curve estimation, *The Annals of Statistics* **21**: 179–195.

Meyer, S., Elias, J. and Höhle, M. (2012). A space–time conditional intensity model for invasive meningococcal disease occurrence, *Biometrics* **68**(2): 607–616.

Møller, J. (2003). Shot noise Cox processes, *Advances in Applied Probability* **35**: 4–26.

Møller, J. and Diaz-Avalos, C. (2010). Structured spatio-temporal shot-noise Cox point process models, with a view to modelling forest fires, *Scandinavian Journal of Statistics* **37**: 2–25.

Møller, J. and Ghorbani, M. (2010). Second-order analysis of structured inhomogeneous spatio-temporal point processes, *Technical Report R-07-2010*, Department of Mathematical Sciences, Aalborg University. Submitted for journal publication.

Møller, J. and Ghorbani, M. (2012). Aspects of second-order analysis of structured inhomogeneous spatio-temporal point processes, *Statistica Neerlandica* **66**: 472–491.

Møller, J. and Ghorbani, M. (2013). Functional summary statistics for the johnson-mehl model, *Technical Report R-01-2013*, Department of Mathematical Sciences, Aalborg University.

Møller, J. and Rasmussen, J. G. (2012). A sequential point process model and bayesian inference for spatial point patterns with linear structures, *Scandinavian Journal of Statistics* **39**(4): 618–634.

Møller, J., Syversveen, A. R. and Waagepetersen, R. P. (1998). Log Gaussian Cox processes, *Scandinavian Journal of Statistics* **25**: 451–482.

Møller, J. and Toftaker, H. (2012). Geometric anisotropic spatial point pattern analysis and cox processes, *Technical report*, Department of Mathematical Sciences, Aalborg University.

Møller, J. and Waagepetersen, R. P. (2004). *Statistical Inference and Simulation for Spatial Point Processes*, Chapman and Hall/CRC, Boca Raton.

Moraga, P. and Montes, F. (2011). Detection of spatial disease clusters with lisa functions, *Statistics in medicine* **30**(10): 1057–1071.

Mugglestone, M. A. and Renshaw, E. (1998). Detection of geological lineations on aerial photographs using two-dimensional spectral analysis, *Computers & Geosciences* **24**(8): 771–784.

Müller, H. G. (1985). On the number of sign changes of a real function, *Periodica Mathematica Hungarica* **16**: 209–213.

Müller, H.-G. and Gasser, T. (1979). Optimal convergence properties of kernel estimates of derivatives of a density function, *in* T. Gasser and M. Rosenblatt (eds), *Smoothing Techniques for Curve Estimation*, Vol. 757 of *Lecture Notes in Mathematics*, Springer Berlin Heidelberg, pp. 144–154.

Nadaraya, E. A. (1964). On estimating regression, *Theory of Probability & Its Applications* **9**: 141–142.

Neyman, J. and Scott, E. L. (1958). Statistical approach to problems of cosmology, *Journal of the Royal Statistical Society Series B* **20**: 1–43.

Ogata, Y. (1998). Space-time point-process models for earthquake occurences, *Annals of the Institute of Statistical Mathematics* **50**: 379–402.

Ohser, J. (1983). On estimators for the reduced second moment measure of point processes, *Mathematische Operationsforschung und Statistik, series Statistics* **14**: 63–71.

Ohser, J. (1991). On estimation of pair correlation functions, *Res Informatics* **4**: 147–152.

Ohser, J. and Mücklich, F. (2000). *Statistical Analysis of Microstructures in Materials Science*, Wiley, New York.

Ohser, J. and Stoyan, D. (1981). On the second-order and orientation analysis of planar stationary point processes, *Biometrical Journal* **23**: 523–533.

Parzen, E. (1962). On estimation of a probability density function and mode, *The Annals of Mathematical Statistics* **33**: 1065–1076.

Peebles, P. J. E. and Groth, E. J. (1975). Statistical analysis of extragalactic objects. V: three-point correlation function for the galaxy distribution in the Zwicky catalog, *Astrophysical Journal* **196**: 1–11.

Perry, J., Liebhold, A., Rosenberg, M., Dungan, J., Miriti, M., Jakomulska, A. and Citron-Pousty, S. (2002). Illustrations and guidelines for selecting statistical methods for quantifying spatial pattern in ecological data, *Ecography* **25**(5): 578–600.

Pielou, E. C. (1977). *Mathematical Ecology*, Wiley-Interscience publication, John Wiley & Sons, New York.

Rathbun, S. L. and Cressie, N. (1994). Space-time survival point processes: longleaf pines in southern Georgia, *Journal of the American Statistical Association* **89**: 1164–1174.

Redenbach, C., Särkkä, A., Freitag, J. and Schladitz, K. (2009). Anisotropy analysis of pressed point processes, *AStA Advances in Statistical Analysis* **93**(3): 237–261.

Ripley, B. D. (1976). The second-order analysis of stationary point processes, *Journal of Applied Probability* **13**: 255–266.

Ripley, B. D. (1977). Modelling spatial patterns (with discussion), *Journal of the Royal Statistical Society Series B* **39**: 172–212.

Ripley, B. D. (1981). *Spatial Statistics*, Wiley, New York.

Ripley, B. D. (1988). *Statistical Inference for Spatial Processes*, Cambridge University Press, Cambridge.

Ripley, B. D. (1989). Gibbsian interaction models, *in* D. A. Griffiths (ed.), *Spatial Statistics: Past, Present and Future*, Image, New York, pp. 1–19.

Rodríguez-Cortés, F. J., Ghorbani, M., Mateu, J. and Stoyan, D. (2014). Variance for estimators of spatio-temporal product density function, *Journal* **00**: 0000–0000.

Rosenberg, M. S. (2004). Wavelet analysis for detecting anisotropy in point patterns, *Journal of Vegetation Science* **15**(2): 277–284.

Rosenblatt, M. (1971). Curve estimates, *The Annals of Mathematical Statistics* **42**: 1815–1842.

Sacks, J. and Ylvisaker, D. (1981). Asymptotically optimum kernels for density estimation at a point, *The Annals of Statistics* **9**: 334–346.

Schladitz, K. and Baddeley, A. (2000). A third-order point process characteristic, *Scandinavian Journal of Statistics* **27**: 657–671.

Schoenberg, F. P., Brillinger, D. R. and Guttorp, P. (2006). *Point Processes, Spatial-Temporal*, John Wiley & Sons, Ltd, Chichester, UK, pp. 2885–2886.

Shumaker, B. P. and Sinnott, R. W. (1984). Astronomical computing: 1. computing under the open sky. 2. virtues of the haversine., *Sky and telescope* **68**: 158–159.

Silverman, B. W. (1978). Distances on circles, toruses and spheres, *Journal of Applied Probability* **15**: 136–143.

Stoyan, D. (1987). Statistical analysis of spatial point processes: A soft-core model and cross-correlations of marks, *Biometrical Journal* **29**: 971–980.

Stoyan, D. and Beneš, V. (1991). Anisotropy analysis for particle systems, *Journal of Microscopy* **164**(2): 159–168.

Stoyan, D., Bertram, U. and Wendrock, H. (1993a). Estimation variances for estimators of product densities and pair correlation functions of planar point processes, *Annals of the Institute of Statistical Mathematics* **45**: 211–221.

Stoyan, D., Bertram, U. and Wendrock, H. (1993b). Estimation variances for estimators of product densities and pair correlation functions of planar point processes, *Annals of the Institute of Statistical Mathematics* **45**: 211–221.

Stoyan, D., Kendall, W. S. and Mecke, J. (1987). *Stochastic Geometry and Its Applications*, Vol. 8, 2nd edn, Wiley Chichester.

Stoyan, D., Kendall, W. S. and Mecke, J. (1995). *Stochastic Geometry and Its Applications*, second edn, Wiley, Chichester.

Stoyan, D. and Stoyan, H. (1994). *Fractals, Random Shapes and Point Fields*, Wiley, Chichester.

Stoyan, D. and Stoyan, H. (2000). Improving ratio estimators of second order point process characteristics, *Scandinavian Journal of Statistics* **27**: 641–656.

Tamayo-Uria, I., Mateu, J. and Diggle, P. (2014). Modelling of the spatio-temporal distribution of rat sightings in an urban environm, *Submitted*
.

Vincenty, T. (1975). Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations, *Survey review* **23**(176): 88–93.

Wand, M. and Ripley, B. (2013). Functions for kernel smoothing for wand and jones (1995). R package version 2.23-10.

Watson, G. S. (1964). Smooth regression analysis, *Sankhyā* **26**: 359–372.

Watson, G. S. and Leadbetter, M. R. (1963). On the estimation of the probability density, i, *Ann. Math. Statist.* **34**: 480–491.

Whittaker, E. and Watson, G. (1996). *A Course of Modern Analysis*, Cambridge University Press.

Woolhouse, W. S. B. (1870). Explanation of a new method of adjusting mortality tables; with some observations upon mr, makeham's modification of gompertz's theory, *Journal of the Institute of Actuaries and Assurance Magazine* **15**: 389 – 410.